


# 随机删失模型中的渐近理论

# Asymptotic Theory for Random Censorship Model

王启华



高等教育出版社
HIGHER EDUCATION PRESS




# 随机删失模型中的渐近理论

# Asymptotic Theory for Random Censorship Model

王启华



高等教育出版社

HIGHER EDUCATION PRESS

图书在版编目（CIP）数据

随机删失模型中的渐近理论 / 王启华. —北京：高等教育出版社，2002.6

统计专业研究生教材

ISBN 7-04-010470-9

Ⅰ.随… Ⅱ.王… Ⅲ.渐近统计量-研究生-教材 Ⅳ.0212

中国版本图书馆 CIP 数据核字（2001）第 091894 号

随机删失模型中的渐近理论

王启华

| | | | |
|---|---|---|---|
| 出版发行 | 高等教育出版社 | 购书热线 | 010－64054588 |
| 社　　址 | 北京市东城区沙滩后街 55 号 | 免费咨询 | 800－810－0598 |
| 邮政编码 | 100009 | 网　　址 | http://www.hep.edu.cn |
| 传　　真 | 010－64014048 | | http://www.hep.com.cn |

| | | | |
|---|---|---|---|
| 经　　销 | 新华书店北京发行所 | | |
| 印　　刷 | 高等教育出版社印刷厂 | | |

| | | | |
|---|---|---|---|
| 开　　本 | 850×1168　1/32 | 版　　次 | 2002 年 6 月第 1 版 |
| 印　　张 | 5.75 | 印　　次 | 2002 年 6 月第 1 次印刷 |
| 字　　数 | 140 000 | 定　　价 | 9.80 元 |
| 插　　页 | 1 | | |

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

# 内 容 提 要


本文第一章第 1.1 节使用 Kaplan-Meier 估计渐近方差的刀切估计，建立了学生化 Kaplan-Meier 估计的 Edgeworth 展开与被估计的 Edgeworth 展开. 证明了两种 Edgeworth 展开逼近其分布的渐近速度均为 $o(n^{\frac{1}{2}})$.

第一章第 1.2 节建立了 Tsai, Jewel 与 Wang[79] 所提出的左截断右删失乘积限估计的 $U$ 统计量表示，利用该表示建立了估计的 Berry-Essen 不等式，而且利用 Gijbels 与 Wang[37] 的定理 1(c)得到了估计的 $r(\geqslant 2)$阶绝对矩不等式与 strassen 型重对数律.

设 $S(s,t)$是二元生存函数，$\hat{S}_n(s,t)$是 Campbell 与 Földes[15] 所提出的二元乘积限估计，第一章第 1.3 节通过将 $\log\hat{S}_n(s,t)-\log S(s,t)$表成 $U$ 统计量加上具有所需性质的余项，利用 $U$ 统计量的 Berry－Essen 定理，建立了二元乘积限估计的 Berry-Essen 不等式.

第二章第 2.1 节给出了随机删失下概率密度估计的 $L_1$ 矩不等式，$p(\geqslant 2)$阶绝对矩不等式，与一个概率不等式，建立了概率密度的光滑 bootstrap 逼近定理，并证明了概率密度 bootstrap 估计的方差几乎处处收敛到概率密度估计的渐近方差.

第二章第 2.2 节构造了失效率的一种核估计，并研究了该估计的弱收敛速度、一致强相合性收敛速度、渐近表示与渐近正态性等问题.

第二章第 2.3 节定义了生存分布的一种平均型泛函的估计，利用点过程理论与鞅方法证明了该估计的渐近正态性，此外，还给出了一个均方误差不等式和一个概率不等式.

第三章研究了随机删失非参数回归模型与半参数回归模型.

在第 3.1 节,我们分别就删失分布已知与未知两种情况构造了非参数回归模型中回归函数的一种加权核估计,并研究了它的一些收敛性质.

对固定设计半参数回归模型

$$Y_i = x_i\beta + g(t_i) + \epsilon_i, \quad i = 1,2,\cdots,n,$$

第 3.2 节也是分别就分布已知与未知两种情况,定义了参数 $\beta$ 与回归函数 $g(\cdot)$ 的估计,证明了它们均具有强相合性与 $p(\geqslant 2)$ 阶平均相合性.

**关键词** 随机删失 乘积限估计 删失回归

# ASYMPTOTIC THEORY FOR RANDOM CENSORSHIP MODEL


## Abstract

In Section 1. 1 of Chapter 1, the jackknife estimte of the asymptoticvariance ofKaplan-Meier (KM) estimator is employed, and the Edgeworth expansions for the studentized Kaplan-Meier estimatorare established. It is shown that the Edgeworth expansions are asympototically close to the exact distribution of the KM estimator with remainder $o(n^{\frac{1}{2}})$ under some mild conditions.

In Section 1.2 of Chapter 1, it is shown that the product-limit (PL) estimator from left truncated and right censored data, which is defined by Tsai, Jewell and Wang[79] may be approximated within a sufficient degree of accuracy by a $U$-statistic. In this way, Berry-Essen theorem for $U$-statistic does carry over to the PL estimator. Moreover, a direct application of Theorem 1 (c) of Gijbels and Wang[37] yields the rth($r>1$)order absolute moment inequality and functional law of iterated logarithm law for the error of the PL estimator.

Let $S(s,t)$ be the bivariate survival function. Let $\hat{S}_n(s,t)$ be the bivariate PL estimator proposed by Campbell and Földes[15]. In Section 1.3 of Chapter 1, the Berry-Essen inequality for $\hat{S}_n(s,t)$ is established by expressing $\hat{S}_n(s,t) - \log S(s,t)$ as a $U$-statistic, which admids a term Edgeworth expansion, plus some remainders with sufficient accuracy.

In Section 2.1 of Chapter 2, the kernel density estimate $\hat{f}_n(t)$ from random censored data is investigated further. $L_1$-moment, $p$ ($\geqslant 2$) order absolute moment inequalities and a probability inequality for the deviation of $\hat{f}_n(t)$ are obtained, respectively. The sufficient conditions for thesmoothed bootstrap approximation of $\hat{f}_n(t)$ to be valid are established. Moreover, it is shown that the variance of the smoothed bootstrap estimator $\hat{f}_n^*(t)$ converges to the asympototic variance of $\hat{f}_n(t)$ almost surely.

In Section 2.2 of Chapter 2, a nonparametric hazard estimator is introduced. Weak conver-gence, strong uniformly consistency of the proposed estimator $\lambda_n(t)$ are investigated on a bounded intervals, respectively. An asymptiotic representation is also given, and the asympototic represen-tation is used to prove asymptotic normality of the hazard estimator.

In Section 2.3 of Chapter 2, a class of mean type functionals of KM estimator are investigated. Counting process martingal methods are used to show the asympototic normality, and establish a mean square error inequality and a probability inequality of them without the assumption that $F$, $G$ are continuous, where $F$, $G$ are survival time distribution and censoring time distribution respectively.

In Section 3.1 of Chapter 3, weighted kernel estimators of regression function in the fixed design nonparametric model are constructed, respectively, in the two cases where censoring distribution is known or unknown. And some convergent properties of them are also investigated under diffrernt conditions.

For the fixed design semiparametric regression model

$$Y_i = x_i\beta + g(t_i) + \epsilon_i, \quad 1 = 1,2,\cdots,n,$$

Section 3.2 of Chapter 3 defines the estimators of unknown parametric $\beta$ and regression function $g(\cdot)$, when the observations

are randomly censored on the right in the two cases: Censoring distribution is known or unknown, respectively. The sufficient conditions under which these estimators are strong consistent and pth( $p \geqslant 2$ ) mean consistent are also established.

**Key Words** Random censorship Product-Limit Estimator Censoring Regression

# 目　录

# 引　言

随机删失模型是生存分析、医药追踪研究、可靠性寿命试验及其他一些实际问题中常用和碰到的一类重要的统计模型. Rupert与Milller[68],郑祖康[108],Fleming与Harrington[31]等已举出很多属于随机删失模型的实例,说明了这一模型存在的广泛性.实事上,随机删失模型的引进使得一些问题的研究更接近现实,这是因为在一些实 际问题中(如以上文献所举实例),我们所获的一些数据并非完全数据,如果为了处理简便而近似地把这些数据看作真实数据,用通常的统计方法去研究解决问题,其结果与现实必然有很大误差,而在随机删失模型下用随机删失模型中发展起来的理论和方法去研究 因随机删失而产生的删失数据就避免了研究的人为性.此外,随机删失模型具有一般性,因为任何一个通常的统计问题都可以看成是相应的随机删失模型下的统计问题在删失发生在无穷远的特殊情形.正如郑祖康所指出:每一个统计问题都可出现一个相应的删失数据的统计问题,当删失不复存在时(即认为删失发生在无穷远)便退化成原来的统计模型.

从以上几点,我们足以看到研究、发展随机删失模型下的统计方法和理论是极其重要的.正因如此,近一二十年来这一领域的研究已异常活跃,特别是非参数估计方面已取得很大进展,如:

## 1.乘积限估计

设 $X_1,X_2,\cdots,X_n$ 是非负独立同分布表示寿命的随机变量序列,具有共同分布函数 $F$.设 $Y_1,Y_2,\cdots,Y_n$是非负独立同分布表示删失的随机变量序列,具有分布函数 $G$,且诸 $X_i$ 独立于诸 $Y_i$,在随机删失模型中,$X_i$ 因被 $Y_i$ 右删失而不能被完全观察,我们仅能观察到

$$Z_i = \min(X_i, Y_i), \quad \delta_i = I[X_i \leqslant Y_i], i = 1,2,\cdots,n.$$

基于上面的删失数据$(Z_i,\delta_i)$,$i = 1,2,\cdots,n$,Kaplan 与 Meier[46] 提出了著名的乘积限估计(或称 Kaplan-Meier 估计,以下简称 KM 估计来估计 $F$,其估计定义如下:

$$1 - \hat{F}_n(t) = \begin{cases} \prod\limits_{i:\, Z_i \leqslant t} \left( \dfrac{N^+(Z_i)}{1 + N^+(Z_i)} \right)^{\delta_i}, & \text{若 } t \leqslant Z_{(n)}, \\ 0, & \text{若 } t > Z_n, \end{cases} \tag{0.1}$$

$$N^+(\cdot) = \sum_{i=1}^{n} I[Z_i > \cdot], \quad Z_{(n)} = \max_{1 \leqslant i \leqslant n} Z_i.$$

自 KM 估计提出十几年后,Breslow 与 Crowley[9] 首次对这一估计进行了研究,并证明了它的弱收敛性.又过了将近十年,Gill[39] 将弱收敛区间扩展到整个半直线.后来,Burke,Csörgo 与 Horváth[11],[12] 又给出了其强逼近的速度,Major 与 Rejtö[54] 还给出了 KM 估计的一个嵌入定理.而对 KM 估计强相合性及其收敛速度的研究更多,主要结果可见 Peterson[59],Winter,Földes 与 Rejtö[86],Földes,Rejtö 与 Winter[34],Földes 与 Rejtö[32],[33],Horváth[42],Gill[39],Csörgo 与 Horváth[23],Shorack 与 Wellner[71],Zheng[94],Wang[84],Gu[40] 等.Lo 与 Singh[51] 将 KM 估计渐近表示成独立同分布平均和加上一余项,并发展了相应的 Bootstrap 理论.Chang 与 Rao[16] 又将 KM 估计表成 $U$ 统计量加上一些具有所需性

质的余项,使得 $U$ 统计量的 Berry-Essen 不等式应用到此表示可获得 KM 估计的 Berry-Essen 不等式. Chang[18] 又利用类似的表示建立了 KM 估计的 Edgeworth 展开.但由于 Chang 等上面的结果是关于标准化 KM 估计建立的,而当 KM 估计渐近方差未知时,他们的结果并不能直接应用到区间估计与假设检验等问题,于是估计 KM 估计的渐近方差并研究学生化 KM 估计的 Edgeworth 展开等问题就成为必然.

本文第一章第 1.1 节使用 KM 估计渐近方差的刀切估计,在 $F,G$ 连续的条件下 建立了学生化 KM 估计的 Edgeworth 展开.

在另一些实际问题中,如天文学、石油勘探和生物细胞的研究中,还会出现这样一种情况,即一些被感兴趣的随机变量无法观察,这主要表现在一些个体在研究开始之前就已经寿终或个体本身因条件限止根本无法观察,从而造成左截断(left truncation).这些实例可见 Ameiya[1], Lynden-Bell[53], Jackson[43], Nicall 与 Segal[58] 与 Rupert 与 Miller[68] 等人的文章和著作,而对那些已进入我们研究的个体,又可能招致如前所述的随机右删失,这种左截断与右删失都发生的情况下所获数据为左截断右删失数据,利用这种数据 Tsai, Jewell 与 Wang[79] 定义了生存函数的乘积限估计, Gijbels 与 Wang[37] 给出了该估计的独立同分布随机变量平均和的强表示定理,同时将所获的强表示结果应用到概率密度和失效率估计的研究.

本文第一章第 1.2 节建立了 Tsai 等人的乘积限估计的 $U$ 统计量表示,利用该表示建立了该估计的 Berry-Essen 不等式, $r(\geqslant 2)$ 阶绝对矩不等式与 Strassen 型重对数律.

关于二元生存分布函数 $S(s,t)=P(X>s,Y>t)$ 的估计最主要有 Campbell[14] 所提出的分层抽样估计与自相容估计, Campbell 与 Földes[15] 所提出的二元乘积限估计, Dabrowska[25] 所提出的二元 Kaplan-Meier 估计.关于二元乘积限估计的研究, Campbell 与 Földes[15] 证明了它的弱收敛性,并给出了它的强一致

性重对数率收敛速度,接着,Wang[85] 证明了它的弱收敛速度与二元经验过程弱收敛速度相同.后来,Lo 与 Wang([52]) 给出了二元乘积限估计的独立同分布渐近表示,由此表示虽可得到它的渐近正态性和重对数率等结果,但若想得到估计的 Berry-Essen 界,这种表示就无能为力了.

本文第一章第 1.3 节将二元乘积限估计的对数表成 $U$ 统计量加上具有所需性质的余项,以至 $U$ 统计量的 Berry-Essen 定理能被应用到该表示以获二元乘积限估计的 Berry-Essen 不等式.

## 2. 乘积限估计的泛函

常见的生存分布函数泛函的估计有概率密度估计、失效率估计与平均生存时间等.

在概率密度估计方面,Blum 与 Susarla[7] 首先提出概率密度核估计,Födels,Rejtö 与 Winter[35],Mielnizuk[56],Wang[84] 等分别研究了概率密度核估计的强相合性的一致收敛速度,$L_1$ 相合性,渐近表示及 $L_p$ 模下的中心极限定理等问题.关于概率密度的其它形式的估计可见 Birgé[6],Zheng[92] 与洪圣岩[100] 等.

本文第二章第 2.1 节给出了概率密度的 $L_1$ 矩不等式,$p(\geqslant 2)$ 阶绝对矩不等式与一个概率不等式,建立了概率密度的光滑 bootstrap 逼近定理,并证明了概率密度 bootstrap 估计的方差几乎处处收敛到概率密度估计的渐近方差.

在失效率估计方面,人们根据失效率的各种不同的表示,给出了它的各种形式的估计并研究了它们的相应的大样本性质,具体的结果可见 Blum 与 Susarla[7],Tanner 与 Wong[77],Tanner[78],Liu 与 Ryzin[50],Sabine 与 Stute[69] 等.

本文第二章第 2.2 节根据失效率的表达式 $\lambda(t) = -\frac{\mathrm{d}}{\mathrm{d}t}\log(1-F(t))$ 构造了它的一种核估计,并研究了该估计的弱收敛速度、一

致强相合性收敛速度、渐近表示与渐近正态性等问题.

又正如 Gijbels 与 Veraverbeke[36] 所指出,许多生存时间分布的特征可表示成

$$\xi_{um} = \int_0^\infty \cdots \int_0^\infty \theta_u(t_1, t_2, \cdots, t_m) \prod_{i=1}^m \mathrm{d}F(t_i), \qquad (0.2)$$

其中 $\theta_u(t_1, t_2, \cdots, t_m)$ 是可能依赖于某参数 $u$ 的 $m$ 元函数.对任何固定的使得 $(1 - F(t))(1 - G(t)) > 0$ 的 $T > 0$,Gijbels 与 Veraverbeke[36] 考虑了

$$\tilde{\xi}_u(F) = \int_0^T \cdots \int_0^T \theta_u(t_1, t_2, \cdots, t_m) \prod_{i=1}^m \mathrm{d}F(t_i)$$

的估计问题,他们以 $F$ 的 KM 估计 $\hat{F}_n$ 取代 $\tilde{\xi}_n(F)$ 中的 $F$,即获得 $\tilde{\xi}_n(F)$ 的估计 $\tilde{\xi}_n(\hat{F}_n)$,并通过将 $\tilde{\xi}_n(\hat{F}_n)$ 表示成 $U$ 统计量,证明了它的渐近正态性,且研究了几乎处处大样本行为.易见 $\tilde{\theta}_u(F)$ 是 (0.2) 在$[0, T]^m$ 上的限制,从实际意义看,这种限制并不是我们所希望的,因而,定义(0.2) 的估计并研究其性质是重要的也是一个困难的问题.

本文第二章第 2.3 节研究了 $\xi_u(F) = \int_0^\infty (1 - F(t))\mathrm{d}\theta_u(t)$ 的估计问题.易见,在一定条件下,(0.2) 在 $m = 1$ 时,即 $\xi_{u1}(F)$,可表示成 $\xi_u(F) + \theta(0)$ 的形式,因此,我们研究 $\xi_u(F)$ 是很有意义的.当我们以 $\hat{F}_n$ 取代$\xi_u(F)$ 中的 $F$ 后即得$\xi_u(F)$ 的估计 $\xi_u(\hat{F}_n)$. 本文第 2.3 节利用点过程鞅方法证明了 $\xi_u(\hat{F}_n)$ 的渐近正态性,并在适当的条件下给出了 $\xi(\hat{F}_n)$ 的一个均方误差不等式和一个概率不等式.

## 3.删失回归模型

关于回归模型研究较多的是删失参数回归模型,这方面结果已很丰富,具体结果可见 Buckley 与 James[10],Koul,Susarla 与

Ryzin[47], Miller 与 Halpern[57], James 与 Smith[44], Zheng[90],[91], Tsiatis[80], Lai 与 Ying[49], Zhou[96], Ying[89].

本文第三章研究了随机删失非参数回归模型与半参数回归模型.

在第 3.1 节我们分别就删失分布已知与未知两种情形构造了非参数回归模型中非参数回归函数的一种加权核估计,并研究了它的一些收敛性质,而强相合性是其中一个概率级数收敛性的简单推论.

在第 3.2 节也是分别就删失分布已知与未知两种情形定义了半参数回归模型中参数 $\beta$ 与回归函数 $g(\cdot)$ 的估计,证明了它们均具有强相合性与 $p(\geqslant 2)$ 阶平均收敛性.

# 第 1 章

# 乘积限估计

本章主要研究由 Kaplan-Meier[45] 所提出的乘积限估计(也称 Kaplan-Meier 估计,以下简称 KM 估计),Tsai,Jewell 与 Wang[79] 所定义的基于左截断右删失的乘积限估计与 Campbell 与 Födles[15] 所构造的二元乘积限估计.

## §1.1 学生化 KM 估计的 Edgeworth 展开

### 1.1.1 记号与主要结果

设 $T_1, T_2, \cdots, T_n$ 是非负独立同分布一般表示寿命的随机变量,有共同的分布函数 $F$. 在随机右删失模型中,每个 $T_i$ 因受另一非负且假定与之独立的随机变量(以后称之为删失随机变量)$C_i$ 的干扰或影响,使得 $T_1, T_2, \cdots, T_n$ 不能被完全观察,而仅能观察到 $(Z_i, \delta_i)$,此处 $Z_i = \min(T_i, C_i)$,$\delta_i = I[T_i \leqslant C_i]$,$i = 1, 2, \cdots, n$,$I[\cdot]$ 表示某事件的示性函数. 本文假定 $C_1, C_2, \cdots, C_n$ 独立同分布有共同的分布函数 $G$. 易见 $Z_1, Z_2, \cdots, Z_n$ 独立同分布且具有共同的分布函数 $H = 1 - (1 - F)(1 - G)$. 设 $Z_{(1)} \leqslant Z_{(2)} \leqslant \cdots \leqslant Z_{(n)}$

是 $Z_1,Z_2,\cdots,Z_n$ 的次序统计量，$\delta_{(i)}$ 是对应于 $Z_{(i)}$ 的 $\delta$.基于上面的删失数据生存函数 $\bar{F}(t)=1-F(t)$ 的 KM 估计定义如下：

$$\widehat{\bar{F}}_n(t)=\prod_{i:Z_{(i)}\leqslant t}\left(\frac{n-i}{n-i+1}\right)^{\delta_{(i)}} I[Z_{(n)}>t]. \quad (1.1.1)$$

对任何分布函数 $Q$，我们定义 $\bar{Q}=1-Q, Q^{-1}=\frac{1}{Q}$，于是 $\bar{H}=\bar{F}\bar{G}$. 记 $\bar{H}_n=n^{-1}\sum_{i=1}^{n}I[Z_i>t]$，$\widetilde{H}_1=P(Z_1>t,\delta_1=1)$，$\widetilde{H}_{n1}(t)=n^{-1}\sum_{i=1}^{n}I[Z_i>t,\delta_i=1]$，$\widetilde{H}_2(t)=P(Z_1>t,\delta_1=0)$，$H_{n2}=n^{-1}\sum_{i=1}^{n}I[Z_i>t,\delta_i=0]$，$\sigma_0^2=\bar{F}^{-2}\sigma^2$，$\sigma^2=-\bar{F}^2(t)\int_0^t\bar{H}^{-2}\mathrm{d}\widetilde{H}_1$.

由 Chang 与 Rao[16]，$\sigma^2$ 是 KM 估计的渐近方差，对此渐近方差 Singh 与 Liu[73] 构造了它的刀切估计.

设 $\widehat{\bar{F}}_n^{(-i)}$ 是基于 $\{(Z_1,\delta_1),\cdots,(Z_n,\delta_n)\}-\{(Z_i,\delta_i)\}$ 的KM估计，其虚拟值随机过程为

$$J_{ni}=n\widehat{\bar{F}}_n(t)-(n-1)\widehat{\bar{F}}_n^{(-i)}(t),\quad t\geqslant 0,1\leqslant i\leqslant n. \quad (1.1.2)$$

由 Singh 与 Liu[73]，$\sigma^2$ 的刀切估计是

$$\hat{\sigma}_{nJ}^2=n^{-1}\sum_{i=1}^{n}(J_{ni}-\bar{J}_n)^2, \quad (1.1.3)$$

其中 $\bar{J}_n=n^{-1}\sum_{i=1}^{n}J_{ni}$. 我们在下面定理 1.1.1，定理 1.1.2 中分别建立学生化 KM 估计 $\sqrt{n}(\widehat{\bar{F}}_n-\bar{F})/\hat{\sigma}_{nJ}$ 的 Edgeworth 展开和被估计的 Edgeworth 展开.

**定理1.1.1** 设 $F$、$G$ 是支撑集分别为 $[0,\infty)$，$[0,\infty)$ 的连续分布函数，则当 $n$ 充分大时，对任何 $t>0$，我们有

$$\sup_x\left|P\left(\sqrt{n}\hat{\sigma}_{nJ}^{-1}(\widehat{\bar{F}}_n(t)-\bar{F}(t))\leqslant x\right)-K_n(x)\right|=o(n^{-\frac{1}{2}}), \quad (1.1.4)$$

其中

$$K_n(x) = \Phi(x) - \frac{\kappa_3}{6} n^{-\frac{1}{2}} \phi(x)(x^2 + \frac{1}{2}), \qquad (1.1.5)$$

$$\kappa_3 = -2\sigma_0^{-3}\left(\int_0^t \overline{H}^{-3} \mathrm{d}\widetilde{H}_1 + \frac{3}{2}\sigma_0^4\right),$$

$\Phi(x)$ 是标准正态分布函数，$\phi(x)$ 是标准正态密度函数.

易见，当 $F$ 未知时，$\sigma_0$ 因而 $\kappa_3$ 未知. 此时，我们必须利用观察 $(Z_i, \delta_i), i = 1, 2, \cdots, n$ 估计 $\sigma_0^2$ 与 $\kappa_3$. 一个自然的做法是由 $\overline{H}_n$ 与 $\widetilde{H}_{n1}$ 分别取代 $\kappa_3$ 与 $\sigma_0^2$ 中的 $\overline{H}$ 与 $\widetilde{H}_1$，即得 $\kappa_3$ 与 $\sigma_0^2$ 的估计 $\kappa_{3n}$ 与 $\sigma_{0n}^2$，其表达式为：

$$\kappa_{3n} = -2\sigma_{0n}^{-3}\left(\int_0^t \overline{H}_n^{-3} \mathrm{d}\widetilde{H}_{n1} + \frac{3}{2}\sigma_{0n}^4\right),$$

$$\sigma_{0n}^2 = -\int_0^t \overline{H}_n^{-2} \mathrm{d}\widetilde{H}_{n1}.$$

记

$$\widetilde{K}_n(x) = \Phi(x) - \frac{\kappa_{3n}}{6} n^{-\frac{1}{2}} \phi(x)(x^2 + \frac{1}{2}).$$

我们的下面定理将证明定理 1.1.1 中的 $K_n(x)$ 被 $\widetilde{K}_n(x)$ 取代后并不影响其渐近精度.

**定理 1.1.2** 在定理 1.1.1 同样的条件下，对任何 $t > 0$，以概率 1 有

$$\sup_x |P(\sqrt{n}\hat{\sigma}_{nJ}^{-1}(\widehat{\overline{F}}_n(t) - \overline{F}(t)) \leqslant x) - \widetilde{K}_n(x)| = o(n^{-\frac{1}{2}}).$$

### 1.1.2 定理的证明

为方便计，我们特约定 $C$ 可表示任何所需的常数，即使在同一式中出现也可表示不同常数.

正如 Chang[17] 所指出，对任何样本容量 $n$ 以正概率 $\widehat{\overline{F}}_n(t) = 0$，因此 $\log \widehat{\overline{F}}_n(t) = -\infty$. 为克服这一研究上的困难，Chang[17] 将样

本空间 $\Omega$ 分成两部分:$\Omega_0^{(n)}$ 与 $\Omega_1^{(n)} = \Omega - \Omega_0^{(n)}$,其中 $\Omega_0^{(n)}$ 如 Chang[17] 中所定义. Chang[17] 证明了 $P(\Omega_1^{(n)}) = o(n^{-k})$,对任何 $k > 0$ 成立.而对 $\omega \in \Omega_0^{(n)}$, $-\infty < \log\widehat{\overline{F}}_n(t) < +\infty$.这说明我们的研究可集中在子样本空间 $\Omega_0^{(n)}$ 上,所获的 Edgeworth 展开结果在 $\Omega$ 上仍成立.这是因为

$$
\begin{aligned}
&P(\sqrt{n}\hat{\sigma}_{nJ}^{-1}(\hat{F}_n(t) - F(t)) \leqslant x) \\
&= P(\sqrt{n}\hat{\sigma}_{nJ}^{-1}(\hat{F}_n(t) - F(t)) \leqslant x, \Omega_0^{(n)}) + o(n^{-k}),
\end{aligned}
$$

对任何 $k > 0$ 成立.

下面就在 $\Omega_0^{(n)}$ 进行我们的研究.

**引理 1.1.1** 若定理 1.1.1 的条件满足,则对任何 $t > 0$,有

$$
\log\widehat{\overline{F}}_n(t) - \log\overline{F}(t) = n^{-1}\sum_{i=1}^{n}\varphi(Z_i, \delta_i; t) + r_{n1}(t), \tag{1.1.6}
$$

$$
\widehat{\overline{F}}_n(t) - \overline{F}(t) = n^{-1}\sum_{i=1}^{n}\overline{F}(t)\varphi(Z_i, \delta_i; t) + r_{n2}(t), \tag{1.1.7}
$$

其中

$$
\begin{aligned}
&\varphi(Z, \delta; t) \\
&= -\left(\int_0^{z \wedge t}[\overline{H}(s)]^{-2}\mathrm{d}\widetilde{H}_1(s) + \overline{H}^{-1}(z)I[z \leqslant t, \delta = 1]\right)
\end{aligned}
$$

且 $r_{n1}(t), r_{n2}(t)$ 对任使得 $n\varepsilon_n \to \infty$ 的 $0 < \varepsilon_n < 1, k > 0$ 和 $t > 0$ 满足

$$
P(|r_{n1}(t)| > \varepsilon_n) \leqslant C(n\varepsilon_n)^{-k}, \tag{1.1.8}
$$

$$
P(|r_{n2}(t)| > \varepsilon_n) \leqslant C(n\varepsilon_n)^{-k}. \tag{1.1.9}
$$

证明:由 Lo 与 Singh[51] 知,当 $n$ 充分大时

$$
\begin{aligned}
P(|r_{n1}| > \varepsilon_n) \leqslant\, & CP\left(\left|\int_0^t(\overline{H}_n - \overline{H})\overline{H}^{-2}\mathrm{d}(\widetilde{H}_{n1} - \widetilde{H}_1)\right| > \frac{\varepsilon_n}{2}\right) \\
& + P(\sup_{0\leqslant s\leqslant t}|\overline{H}_n - \overline{H}| > \frac{1}{2}H(t))
\end{aligned}
$$

$$+ CP\left(\left|\log\widehat{F}_n(t) - \int_0^t [\overline{H}_n(s)]^{-1} d\widetilde{H}_{n1}(s)\right| > \frac{\varepsilon_n}{2}\right) \quad (1.1.10)$$

对任何 $t > 0$ 与 $\varepsilon_n > 0$ 成立.

易得

$$\begin{aligned}
\Delta_n &\overset{d}{=} \int_0^t (\overline{H}_n - \overline{H})\overline{H}^{-2} d(\widetilde{H}_{n1} - \widetilde{H}_1) \\
&= -\int_0^t \left[\overline{H}^{-2}(s)\int_s^{+\infty} d(\overline{H}_n - \overline{H})(u)\right] d(\widetilde{H}_{n1} - \widetilde{H}_1)(s) \\
&= \int_0^{+\infty}\int_0^{+\infty} (-I[0 \leqslant s \leqslant t, s < u]) \\
&\quad \cdot \overline{H}^{-2}(s) d(\overline{H}_n - \overline{H})(u) d(\widetilde{H}_{n1} - \widetilde{H}_1)(s).
\end{aligned}$$

由 Chang 与 Rao[16] 附录中的引理,对任何正整数 $m$ 和某常数 $C$(可能依赖 $m, \overline{H}, \widetilde{H}_1$)

$$\begin{aligned}
E\Delta_n^m &= E\Big[\int_0^{+\infty}\cdots\int_0^{+\infty}\prod_{i=1}^m (-I[0 \leqslant s_i \leqslant t, s_i < u_i]\overline{H}^{-2}(s_i)) \\
&\quad \times \prod_{j=1}^m d(\overline{H}_n(u_j) - \overline{H}(u_j))\prod_{l=1}^m d(\widetilde{H}_{n1}(s_l) - \widetilde{H}_1(s_l))\Big] \\
&= Cn^{-m},
\end{aligned} \quad (1.1.11)$$

记

$$m = \begin{cases} [k] + 1, & \text{若}[k] + 1 \text{是偶数}, \\ [k] + 2, & \text{若}[k] + 1 \text{是奇数}. \end{cases}$$

由 Tchebyschev 不等式与(1.1.11),且注意到 $m$ 是偶数,我们可得

$$P(|\Delta_n| > \varepsilon_n) \leqslant \varepsilon_n^{-m} E\Delta_n^m \leqslant C(n\varepsilon_n)^{-m} \leqslant C(n\varepsilon_n)^{-k}, \quad (1.1.12)$$

对任何 $t > 0, k > 0$ 与 $\varepsilon_n > 0$ 成立. 又由于

$$\begin{aligned}
\Delta'_n &\overset{d}{=} \log\widehat{F}_n - \int_0^t [\overline{H}_n(s)]^{-1} d\widetilde{H}_{n1}(s) \\
&= \sum_{i=1}^n I[Z_i \leqslant t, \delta_i = 1]\left\{\log\left(1 - \frac{1}{1 + n\overline{H}_n(Z_i)}\right) + \frac{1}{n\overline{H}_n(Z_i)}\right\},
\end{aligned} \quad (1.1.13)$$

记 $A_{n\varepsilon}=\bigcap_{i=1}^{n}\{\{\omega:(1+n\overline{H}_n(Z_i))I[Z_i\leqslant t]\geqslant 2\}\cup\{Z_i>t\}\}$，$A_{n\varepsilon}^c$ 为其补，则由(1.1.13)并利用不等式 $|\log(1-x)+x|\leqslant x^2$，$0\leqslant x\leqslant\frac{1}{2}$，可得在 $A_{n\varepsilon}$ 上，有

$$\Delta'_n\leqslant 2\sum_{i=1}^{n}I[Z_i\leqslant t,\delta_i=1]\frac{1}{n\overline{H}_n(Z_i)(1+n\overline{H}_n(Z_i))}. \tag{1.1.14}$$

而

$$\begin{aligned}P(A_{n\varepsilon}^c)&\leqslant\sum_{i=1}^{n}P((1+n\overline{H}_n(Z_i))I[Z_i\leqslant t]=1)\\&=\sum_{i=1}^{n}\int_0^{T_H}P((1+n\overline{H}_n(Z_i))I[Z_i\leqslant t]=1\mid Z_i=x)\mathrm{d}H(x)\\&\leqslant n\int_0^t\binom{n-1}{0}H^{n-1}(x)\mathrm{d}H(x)\\&\leqslant(1-\overline{H})^n\leqslant \mathrm{e}^{-n\overline{H}(t)}.\end{aligned} \tag{1.1.15}$$

于是由(1.1.14)与(1.1.15)，即得

$$P(|\Delta'_n|>\frac{\varepsilon_n}{2})\leqslant P\left(\frac{1}{n}\sum_{i=1}^{n}I[Z_i\leqslant t,\delta_i=1]>\frac{n\varepsilon_n\overline{H}^2(t)}{16}\right)+P(\overline{H}_n(t)<\frac{1}{2}\overline{H}(t))+P(A_{n\varepsilon}^c). \tag{1.1.16}$$

注意到 $n\varepsilon_n<16\overline{H}^{-2}(t)$ 时(1.1.8)是平凡的，因而只考虑 $n\varepsilon_n\geqslant 16\overline{H}^{-2}(t)$ 的情形，在此情况下(1.1.16)的第一项为0，而

$$\begin{aligned}&P(\overline{H}_n(t)<\frac{1}{2}\overline{H}(t))\\&\leqslant P(|H_n(t)-H(t)|>\frac{1}{2}\overline{H}(t))\leqslant \mathrm{e}^{-n\overline{H}^2(t)}.\end{aligned} \tag{1.1.17}$$

因而由(1.1.15)－(1.1.17)就得到 $n\varepsilon_n\geqslant 16\overline{H}^{-2}(t)$ 时

$$P(|\Delta'_n|>\frac{\varepsilon_n}{2})\leqslant \mathrm{e}^{-n\overline{H}^2(t)}+\mathrm{e}^{-n\overline{H}(t)}. \tag{1.1.18}$$

联合(1.1.10)，(1.1.12)，(1.1.18)与(1.1.7)中的最后一个不等

式，我们就证明了(1.1.8)。

为证(1.1.9)，我们首先证明

$$P(|\log\widehat{\bar F}_n(t)-\log\bar F(t)|>\varepsilon_n)\leqslant C(n^{\frac{1}{2}}\varepsilon_n)^{-k}, \tag{1.1.19}$$

对任何 $t>0, k>0$ 与 $\varepsilon_n>0$ 成立。

由(1.1.6)，我们有

$$P(|\log\widehat{\bar F}_n(t)-\log\bar F(t)|>\varepsilon_n)\leqslant P\left(\left|n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i;t)\right|>\frac{\varepsilon_n}{2}\right)+P\left(|r_{n1}(t)|>\frac{\varepsilon_n}{2}\right). \tag{1.1.20}$$

根据 Tchebyschev 与 Dharmadhikari-Jodgeo (D-J) 不等式(见 Rao[64])，得到

$$\begin{aligned}&P\left(\left|n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i;t)\right|>\frac{\varepsilon_n}{2}\right)\\&\leqslant 2^k(n\varepsilon_n)^{-k}E\left|\sum_{i=1}^n\varphi(Z_i,\delta_i;t)\right|^k\\&\leqslant C(n\varepsilon_n)^{-k}n^{\frac{k}{2}-1}\sum_{i=1}^nE\left|\varphi(Z_i,\delta_i;t)\right|^k\leqslant C(n^{\frac{1}{2}}\varepsilon_n)^{-k}.\end{aligned} \tag{1.1.21}$$

因此(1.1.8)，(1.1.21) 与(1.1.20) 一起就证明了(1.1.19)。

由 Taylor 公式，我们得到

$$\widehat{\bar F}_n-\bar F=\bar F(\log\widehat{\bar F}_n-\log\bar F)+\frac{\bar F e^{\theta(\log\widehat{\bar F}_n-\log\bar F)}}{2}(\log\widehat{\bar F}_n-\log\bar F)^2,\quad 0<\theta<1. \tag{1.1.22}$$

由(1.1.22)，(1.1.6) 与(1.1.7)，易见

$$P(|r_{n2}(t)|>\varepsilon_n)\leqslant P\left(|r_{n1}(t)|>\frac{\varepsilon_n}{2\bar F}\right)$$

$$+ P\left(\frac{\bar{F}(t)\mathrm{e}^{\theta(\log\widehat{\bar{F}}_n - \log\bar{F})}}{2}(\log\widehat{\bar{F}}_n(t) - \log\bar{F}(t))^2 > \frac{\varepsilon_n}{2}\right). \tag{1.1.23}$$

设 $P_n$ 表示(1.1.23)右边第二项.由(1.1.19),我们有

$$\begin{aligned} P_n &\leqslant P(\bar{F}\mathrm{e}^{|\log\widehat{\bar{F}}_n - \log\bar{F}|}(\log\widehat{\bar{F}}_n - \log\bar{F})^2 > \varepsilon_n, |\log\widehat{\bar{F}}_n - \log\bar{F}| \leqslant -\log\bar{F}) \\ &\quad + P(|\log\widehat{\bar{F}}_n - \log\bar{F}| > -\log\bar{F}) \\ &\leqslant P(|\log\widehat{\bar{F}}_n - \log\bar{F}| > \varepsilon_n^{\frac{1}{2}}) + P(|\log\widehat{\bar{F}}_n - \log\bar{F}| > -\log\bar{F}) \\ &\leqslant C(n\varepsilon_n)^{\frac{k}{2}}. \end{aligned} \tag{1.1.24}$$

因此,联合(1.1.23),(1.1.24)与(1.1.8)再注意到 $k$ 的任意性,即证(1.1.9).

为简便计,以下以 $\varphi(z,\delta)$ 记 $\varphi(z,\delta;t)$, $r_{ni}$ 记 $r_{ni}(t)$,此处 $i=1,2,\cdots,n$.

**引理 1.1.2** 在定理1.1.1的假设下对任何 $t>0, k>0$ 与 $\varepsilon_n>0$,我们有

$$\hat{\sigma}_{nJ}^2 = n^{-1}\sum_{i=1}^{n}\bar{F}^2(\varphi(Z_i,\delta_i) - n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i))^2 + R_n(t),$$

而

$$P(|R_n(t)| > \varepsilon_n) \leqslant Cn^{1-\frac{k}{2}}\varepsilon_n^{-k}, \tag{1.1.25}$$

对任何 $t>0, k>0$ 与 $\varepsilon_n>0$ 当 $n$ 充分大时成立.

证明:由(1.1.7)与(1.1.2),得

$$J_{ni} = \bar{F} + \bar{F}\varphi(Z_i,\delta_i) + nr_{n2} - (n-1)r_{n2}^{-i}, \tag{1.1.26}$$

其中 $r_{n2}^{-i}$ 是基于 $\{(Z_1,\delta_1),\cdots,(Z_n,\delta_n)\} - \{(Z_i,\delta_i)\}$ 构造的与 $r_{n2}$ 相应的量.

记

$$\bar{H}_n^{(-i)} = \frac{1}{n-1}\sum_{\substack{j=1\\ j\neq i}}^{n} I[Z_j > t],$$

$$r_{n2i} = nr_{n2} - (n-1)r_{n2}^{(-i)}.$$

由 Singh 与 Liu[73],我们有下面不等式

$$|r_{n2i}| \leqslant r_{n2} + C(|\overline{H}_n^{(-i)}(t) - \overline{H}(t)| + |\widetilde{H}_{n1}(t) - \widetilde{H}(t)|). \tag{1.1.27}$$

因此,根据引理 1.1.1 与 D-J 不等式可得

$$\begin{aligned} & P(\max_{1\leqslant i\leqslant n}|r_{n2i}| > \varepsilon_n) \\ \leqslant & P\left(|r_{n2}| > \frac{\varepsilon_n}{3}\right) + P\left(|\widetilde{H}_{n1} - \widetilde{H}_1| > \frac{C\varepsilon_n}{3}\right) \\ & + \sum_{i=1}^n P\left(|\widetilde{H}_n^{(-i)} - \widetilde{H}| > \frac{C\varepsilon_n}{3}\right) \\ \leqslant & C(n\varepsilon_n)^{-\frac{k}{2}} + C(n^{\frac{1}{2}}\varepsilon_n)^{-k} + Cn((n-1)^{\frac{1}{2}}\varepsilon_n)^{-k} \leqslant Cn^{1-\frac{k}{2}}\varepsilon_n^{-k}. \end{aligned} \tag{1.1.28}$$

由(1.1.3) 与(1.1.26),易见

$$\begin{aligned} \hat{\sigma}_{nJ}^2 = & n^{-1}\sum_{i=1}^n \overline{F}^2\Big(\varphi(Z_i,\delta_i) - n^{-1}\sum_{i=1}^n \varphi(Z_i,\delta_i)\Big)^2 \\ & + 2n^{-1}\sum_{i=1}^n \overline{F}(\varphi(Z_i,\delta_i) - n^{-1}\sum_{i=1}^n \varphi(Z_i,\delta_i))(r_{n2i} - n^{-1}\sum_{i=1}^n r_{n2i}) \\ & + n^{-1}\sum_{i=1}^n \Big(r_{n2i} - n^{-1}\sum_{i=1}^n r_{n2i}\Big)^2. \end{aligned} \tag{1.1.29}$$

设 $V_{n1}, V_{n2}$ 分别定义(1.1.29) 右边第二、第三项,且设 $V_n = V_{n1} + V_{n2}$. 由(1.1.17) 并注意到 $\varphi(z,\delta)$ 是有界随机变量,我们得到

$$\begin{aligned} P(|V_{n1}| > \varepsilon_n) \leqslant & P\left(\left|\frac{2}{n}\sum_{i=1}^n \overline{F}\varphi(Z_i,\delta_i) r_{n2i}\right| > \frac{\varepsilon_n}{2}\right) \\ & + P\left(\left|\frac{2}{n}\sum_{i=1}^n \overline{F}\varphi(Z_i,\delta_i)\frac{1}{n}\sum_{i=1}^n r_{n2i}\right| > \frac{\varepsilon_n}{2}\right) \\ \leqslant & 2P(\max_{1\leqslant i\leqslant n}|r_{n2i}| > C\varepsilon_n) \\ \leqslant & Cn^{1-\frac{k}{2}}\varepsilon_n^{-k}. \end{aligned} \tag{1.1.30}$$

类似地

$$P(|V_{n2}| > \varepsilon_n) \leqslant P\left(n^{-1}\sum_{i=1}^{n} r_{n2i}^2 > \varepsilon_n\right)$$

$$\leqslant P(\max^{1\leqslant i\leqslant n} |r_{n2i}| > \varepsilon_n^{\frac{1}{2}})$$

$$\leqslant Cn^{1-\frac{k}{2}}\varepsilon_n^{-\frac{k}{2}}. \qquad (1.1.31)$$

由(1.1.30)与(1.1.31)对任何 $t > 0, k > 0$ 与 $0 < \varepsilon_n < 1$,有

$$P(|V_n| > \varepsilon_n) \leqslant P(|V_{n1}| > \frac{\varepsilon_n}{2}) + P(|V_{n2}| > \frac{\varepsilon_n}{2})$$

$$\leqslant Cn^{1-\frac{k}{2}}\varepsilon_n^{-k},$$

这就完成了引理 1.1.2 的证明.

**引理 1.1.3** 若定理 1.1.1 的条件满足,则对任何 $t > 0$, $k > 0$ 与 $0 < \varepsilon_n < 1$ 有

$$P(|\hat{\sigma}_{nJ}^2 - \sigma^2| > \varepsilon_n) \leqslant Cn^{1-\frac{k}{2}}\varepsilon_n\}^{-k}.$$

证明:由引理 1.1.2,易见

$$P(|\hat{\sigma}_{nJ}^2 - \sigma^2| > \varepsilon_n)$$

$$\leqslant P\left(\left|n^{-1}\sum_{i=1}^{n}\bar{F}^2\varphi^2(Z_i,\delta_i) - \sigma^2\right| > \frac{\varepsilon_n}{3}\right)$$

$$+ P\left(\left(\frac{1}{n}\sum_{i=1}^{n}\bar{F}\varphi(Z_i,\delta_i)\right)^2 > \frac{\varepsilon_n}{3}\right) + P\left(|R_n| > \frac{\varepsilon_n}{3}\right). \qquad (1.1.32)$$

注意到

$$E\varphi^2(Z,\delta) = \bar{F}^{-2}\sigma^2 \qquad (1.1.33)$$

(见 Lo 与 Singh[51]),且应用证明(1.1.21) 的方法,可证(1.1.32) 右边第一、第二项分别不超过 $C(n^{\frac{1}{2}}\varepsilon_n)^{-k}$ 与 $C(n\varepsilon_n)^{-\frac{k}{2}}$. 再次应用(1.1.25),引理得证.

**引理 1.1.4** 若定理 1.1.1 的条件满足,则对任何 $t > 0$,有

$$\log\widehat{F}_n - \log\bar{F} = U_{n0} - \frac{1}{2}n^{-1}\sigma_0^2 + \Delta_{n1},$$

且对任何 $t > 0$

$$P(\sqrt{n}\,|\Delta_{n1}| > n^{-\frac{1}{2}}(\log n)^{-1}) = o(n^{-\frac{1}{2}}),$$

其中

$$U_{n0} = n^{-2}\delta\sum\sum_{i<j}h_0(Z_i,\delta_i;\ Z_j,\delta_j),$$

$$h_0(Z_1,\delta_1;\ Z_2,\delta_2) = g_0(Z_1,\delta_1) + g_0(Z_2,\delta_2) + \psi_0(Z_1,\delta_1;Z_2,\delta_2),$$

$$g_0(Z,\delta) = -A_{10}(Z)\delta - A_{20}(Z\wedge t),$$

$$A_{10}(s) = \overline{H}^{-1}(s)I[0\leqslant s\leqslant t],$$

$$A_{20}(s) = \int_0^s\overline{H}^{-2}\mathrm{d}\widetilde{H}_1,$$

$$\psi_0(Z_1,\delta_1;\ Z_2,\delta_2) = \eta(Z_1,\delta_1;\ Z_2,\delta_2) + \eta(Z_2,\delta_2;Z_1,\delta_1),$$

$$\eta(Z_1,\delta_1;Z_2,\delta_2) = B_1(Z_1,Z_2)\delta_1 + B_2(Z_1,Z_2) - E[B_1(Z_1,Z_2)\delta_1 + B_2(Z_1,Z_2)\,|\,Z_1,\delta_1],$$

$$B_1(s,u) = \overline{H}^{-2}(s)I[0\leqslant s\leqslant t, s<u],$$

$$B_2(s,u) = \int_0^{t\wedge s\wedge u}\overline{H}^{-3}\mathrm{d}\widetilde{H}_1.$$

证明：见 Chang[17].

注：这里 $h_0(Z_1,\delta_1;\ Z_2,\delta_2)$，$g_0(Z_1,\delta_1)$ 与 $\psi_0(Z_1,\delta_1;Z_2,\delta_2)$ 分别与 Chang[17] 中 $-h(Z_1,\delta_1;Z_2,\delta_2)$，$-g(Z_1,\delta_1)$ 与 $-\psi(Z_1,\delta_1;Z_2,\delta_2)$ 相同，下面，我们将经常使用这一事实.

**引理 1.1.5** 对任何随机变量 $X,Y$，存在常数 $\alpha$，使得对任何常数 $a>0$，有

$$\sup_x|P(X+Y\leqslant x) - K_n(x)| = \sup_x|P(X\leqslant x) - K_n(x)| + \alpha a + P(|Y|>a),$$

其中 $K_n(x)$ 如定理 1.1.1 所定义.

证明：证明与 Chang 与 Rao[16] 中引理 2 的证明相似. 这里略去其细节.

记

$$U_n = n^{-2}\sum\sum_{i<j}h(Z_i,\delta_i;Z_j,\delta_j),\qquad(1.1.34)$$

其中

$$
\begin{aligned}
h(Z_i,\delta_i;\ Z_j,\delta_j)= & h_0(Z_i,\delta_i;\ Z_j,\delta_j) \\
& -\frac{\varphi^2(Z_i,\delta_i)\varphi(Z_j,\delta_j)+\varphi(Z_i,\delta_i)\varphi^2(Z_j,\delta_j)}{2\sigma_0^2} \\
& +\frac{\varphi(Z_i,\delta_i)+\varphi(Z_j,\delta_j)}{2}+\varphi(Z_i,\delta_i)\varphi(Z_j,\delta_j).
\end{aligned}
$$

**引理 1.1.6** 设 $\sigma_n^2$ 是 $U_n$ 的方差,若定理 1.1.1 的条件满足,则当 $n\to+\infty$ 时

$$
\sup_x |P(\sqrt{n}\sigma_n^{-1}U_n\leqslant x)-K_{n0}(x)|=o(n^{-\frac{1}{2}}),
$$

其中

$$
K_{n0}(x)=\Phi(x)-\frac{\kappa_3}{6}n^{-\frac{1}{2}}\phi(x)(x^2-1),
$$

$$
\kappa_3=-2\sigma_0^{-3}\left(\int_0^t \overline{H}^{-3}\mathrm{d}\widetilde{H}_1+\frac{3}{2}\sigma_0^4\right).
$$

证明:我们根据 Chang[17] 的思想证明该引理. 显然 $h(Z_1,\delta_1;Z_2,\delta_2)$ 关于 $(Z_1,\delta_1)$, $(Z_2,\delta_2)$ 对称, i.e., $h(Z_1,\delta_1;\ Z_2,\delta_2)=h(Z_2,\delta_2;Z_1,\delta_1)$. 根据 $E\varphi(Z_1,\delta_1)=0$(见 Lo 与 Wang[52]),

$$
Eh_0(Z_1,\delta_1;\ Z_2,\delta_2)=0
$$

(见 Chang[17]),且注意到 $(Z_1,\delta_1),(Z_2,\delta_2),\cdots,(Z_n,\delta_n)$ 独立同分布,于是我们得到

$$
Eh(Z_1,\delta_1;\ Z_2,\delta_2)=0. \tag{1.1.35}
$$

易见 $U_n$ 是具有对称核 $h$ 的二阶 $U$ 统计量. 设

$$
g(Z_1,\delta_1)\stackrel{\mathrm{d}}{=}E[h(Z_1,\delta_1;Z_2,\delta_2)\,|\,Z_1,\delta_1].
$$

由(1.1.35)与(1.1.33),得

$$
\begin{aligned}
g(Z_1,\delta_1)= & E[h_0(Z_1,\delta_1;Z_2,\delta_2)\,|\,Z_1,\delta_1] \\
& -\frac{\varphi(Z_1,\delta_1)\sigma_0^2}{2\sigma_0^2}+\frac{\varphi(Z_1,\delta_1)}{2} \\
= & g_0(Z_1,\delta_1).
\end{aligned} \tag{1.1.36}
$$

因 $U_n$ 能表示成

$$U_n = n^{-2}\Big[(n-1)\sum_{i=1}^{n} g(Z_i,\delta_i) + \sum\sum_{i<j}\psi(Z_i,\delta_i;Z_j,\delta_j)\Big],$$

其中

$$\psi(Z_1,\delta_1;Z_2,\delta_2) = h(Z_1,\delta_1;\ Z_2,\delta_2) - g(Z_1,\delta_1) - g(Z_2,\delta_2). \tag{1.1.37}$$

易见 $n^2U_n$ 与 BGZ[8] 中(1.5)有相同的形式.现在为证引理1.1.6,只须证

$$\kappa_3 = \sigma_0^{-3}\{Eg^3(Z_1,\delta_1) + 3E[g(Z_1,\delta_1)g(Z_2,\delta_2)\psi(Z_1,\delta_1;Z_2,\delta_2)]\} \tag{1.1.38}$$

与 BGZ[8] 中条件(1.13)与(1.15) – (1.18)满足.

首先证(1.1.38)成立.由(1.1.36)与 Chang[17],我们有

$$Eg^3(Z_1,\delta_1) = Eg_0^3(Z_1,\delta_1) = \int_0^t \overline{H}^{-3}\mathrm{d}\widetilde{H}_1 + \frac{3}{2}\sigma_0^4. \tag{1.1.39}$$

由(1.1.37),(1.1.34)与(1.1.36),我们得

$$\begin{aligned}
&E\{g(Z_1,\delta_1)g(Z_2,\delta_2)\psi(Z_1,\delta_1;\ Z_2,\psi_2)\}\\
&= \{g_0(Z_1,\delta_1)g_0(Z_2,\delta_2)[\psi_0(Z_1,\delta_1;\ Z_2,\delta_2)\\
&\quad - \frac{\varphi^2(Z_1,\delta_1)\varphi(Z_2,\delta_2) + \varphi(Z_1,\delta_1)\varphi^2(Z_2,\delta_2)}{2\sigma_0^2}\\
&\quad + \frac{\varphi(Z_1,\delta_1) + \varphi(Z_2,\delta_2)}{2} + \varphi(Z_1,\delta_1)\varphi(Z_2,\delta_2)]\}\\
&= Eg_0(Z_1,\delta_1)g_0(Z_2,\delta_2)\psi_0(Z_1,\delta_1;Z_2,\delta_2)\\
&\quad - \left\{\frac{2E[g_0(Z_1,\delta_1)\varphi^2(Z_1,\delta_1)]Eg_0(Z_1,\delta_1)\varphi(Z_1,\delta_1)}{2\sigma_0^2}\right.\\
&\quad - E[g_0(Z_1,\delta_1)\varphi(Z_1,\delta_1)]Eg_0(Z_1,\delta_1)\\
&\quad \left. - E^2[g_0(Z_1,\delta_1)\varphi(Z_1,\delta_1)]\right\}
\end{aligned}$$

$$\stackrel{d}{=} E_{1n} - E_{2n}. \tag{1.1.40}$$

由引理 1.1.4 下的注与 Chang[17] 中(6) 的证明,我们有

$$E_{1n} = -\sigma_0^4. \tag{1.1.41}$$

由引理 1.1.1 与引理 1.1.4 中对 $\varphi(Z,\delta)$ 与 $g_0(Z,\delta)$ 的定义,易见

$$g_0(Z_1,\delta_1) = \varphi(Z_1,\delta_1). \tag{1.1.42}$$

因此由(1.1.42),(1.1.39) 及 $E\varphi^2 = \sigma_0^2, E\varphi = 0$ (见 Lo 与 Singh[51]) 这两个结果,我们有

$$Eg_0(Z_1,\delta_1)\varphi^2(Z_1,\delta_1) = Eg_0^3(Z_1,\delta_1) = \int_0^t \overline{H}^{-3}\mathrm{d}\widetilde{H}_1 + \frac{3}{2}\sigma_0^4, \tag{1.1.43}$$

$$Eg_0(Z_1,\delta_1)\varphi(Z_1,\delta_1) = E\varphi^2(Z_1,\delta_1) = \sigma_0^2. \tag{1.1.44}$$

使用(1.1.43) 与(1.1.44) 及 $E\varphi(Z_1,\delta_1) = 0$ 这一结果,即得

$$E_{2n} = \int_0^t \overline{H}^{-3}\mathrm{d}\widetilde{H}_1 + \frac{3}{2}\sigma_0^4 - \sigma_0^4, \tag{1.1.45}$$

联合(1.1.40),(1.1.41) 与(1.1.45),我们得到

$$E\{g(Z_1,\delta_1)g(Z_2,\delta_2)\psi(Z_1,\delta_1;\ Z_2,\delta_2)\}$$
$$= -\left(\int_0^t \overline{H}^{-3}\mathrm{d}\widetilde{H}_1 + \frac{3}{2}\sigma_0^4\right). \tag{1.1.46}$$

因此由(1.1.39),(1.1.46) 与引理中对 $\kappa_3$ 的定义即证得(1.1.38).

下面我们将证明 BGZ[8] 中条件(1.13) 与(1.15) - (1.18) 在此成立.

注意到 $\psi(Z_1,\delta_1;\ Z_2,\delta_2)$ 能表示成

$$\psi(Z_1,\delta_1;\ Z_2,\delta_2) = \psi_0(Z_1,\delta_1;\ Z_2,\delta_2)$$
$$- \frac{\varphi^2(Z_1,\delta_1)\varphi(Z_2,\delta_2) + \varphi(Z_1,\delta_1)\varphi^2(Z_2,\delta_2)}{2\sigma_0^2}$$
$$+ \frac{\varphi(Z_1,\delta_1) + \varphi(Z_2,\delta_2)}{2} + \varphi(Z_1,\delta_1)\varphi(Z_2,\delta_2), \tag{1.1.47}$$

由 Chang[17] 的结果 $E|\psi_0(Z_1,\delta_1;\ Z_2,\delta_2|^r < +\infty, r > 0$ 及 $\varphi(Z,\delta)$ 的有界性,容易得到

$$E\,|\psi(Z_1,\delta_1;\,Z_2,\delta_2)|^r < +\infty, \qquad (1.1.48)$$

对任何 $r > 0$ 成立.

由(1.1.36),(1.1.42)及 $\varphi(z,\delta)$ 的有界性知,存在 $M > 0$ 使得

$$I[\,|g(Z,\delta)| > M] = 0, \qquad (1.1.49)$$

因此,BGZ 中的条件(1.13)与(1.15)-(1.17)满足.

注意到这里的 $-g(Z,\delta)$ 即为 Chang[17] 中的 $g(Z,\delta)$,因而由 Chang[17] 对 BGZ[8] 条件(1.18)的验证结果即知该条件在此亦满足.

**引理 1.1.7** 在引理 1.1.6 的假定下,当 $n$ 充分大时,我们有

$$\sup_x |P(\sqrt{n}\sigma_0^{-1}U_n \leqslant x) - K_{n0}(x)| = o(n^{-\frac{1}{2}}),$$

证明:利用引理 1.1.5,1.1.6,并套用 Chang[17] 中引理 3 的证明即可证引理 1.1.7.

**引理 1.1.8** 在引理 1.1.6 的假定下,当 $n$ 充分大时,我们有

$$\sup_x |P(\sqrt{n}\sigma_0^{-1}U_n - An^{-\frac{1}{2}} \leqslant x) - K_n(x)| = o(n^{-\frac{1}{2}}),$$

其中 $K_n(x)$ 如定理 1.1.1 所定义,

$$A = \frac{1}{2}\sigma_0^{-3}\left(\int_0^t \overline{H}^{-3}\mathrm{d}\widetilde{H}_1 + \frac{3}{2}\sigma_0^4\right).$$

证明:由引理 1.1.6 与 Taylor 公式,我们有

$$\begin{aligned}
&\sup_x |P(\sqrt{n}\sigma_0^{-1}U_n - An^{-\frac{1}{2}} \leqslant x) - K_n(x)| \\
&= \sup_x |P(\sqrt{n}\sigma_0^{-1}U_n \leqslant x + An^{-\frac{1}{2}}) - K_n(x)| \\
&\leqslant \sup_x |P(\sqrt{n}\sigma_0^{-1}U_n \leqslant x) - K_{n0}(x)| \\
&\quad + \sup_x |K_{n0}(x + An^{-\frac{1}{2}}) - K_n(x)| \\
&= \sup_x \left|K_{n0}(x) + \left[\frac{\mathrm{d}}{\mathrm{d}x}K_{n0}(x)\right]An^{-\frac{1}{2}} - K_n(x)\right| + o(n^{-\frac{1}{2}}) \\
&= o(n^{-\frac{1}{2}}).
\end{aligned}$$

于是引理 1.1.8 得证.

定理 1.1.1 的证明：下面我们固定 $t$ 使得 $t > 0$. 令 $Q(x) = \frac{1}{1+x} - (1-x)$，则

$$
\begin{aligned}
&\sqrt{n}\hat{\sigma}_{nJ}^{-1}(\widehat{F}_n - \overline{F}) \\
&= \sqrt{n}\sigma^{-1}(\widehat{F}_n - \overline{F})\left[1 - \frac{\hat{\sigma}_{nJ} - \sigma}{\sigma} + Q\left(\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right)\right] \\
&= \sqrt{n}\sigma^{-1}(\widehat{F}_n - \overline{F})\left(1 - \frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right) \\
&\quad + \sqrt{n}\sigma^{-1}(\widehat{F}_n - \overline{F})Q\left(\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right) \\
&\stackrel{\mathrm{d}}{=} R_{1n} + R_{2n}.
\end{aligned} \tag{1.1.50}
$$

易见

$$
\begin{aligned}
&P(|R_{2n}| > (n\log n)^{-\frac{1}{2}}) \\
&\leqslant P(|\sqrt{n}\sigma^{-1}(\widehat{F}_n - \overline{F}| > \log^{\frac{1}{2}} n) \\
&\quad + P\left(\left|\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right)\right| > n^{-\frac{1}{2}}\log^{-1} n) \\
&\stackrel{\mathrm{d}}{=} R_{21n} + R_{22n}.
\end{aligned} \tag{1.1.51}
$$

记

$$
\Psi_n(x) = \Phi(x) - \frac{\tilde{\kappa}_3}{6} n^{-\frac{1}{2}} \phi(x)(x^2 - 1),
$$

其中

$$
\tilde{\kappa}_3 = -\sigma_0^{-3}\left(-\int_0^t \overline{H}^{-3}\mathrm{d}\widetilde{H}_1 + \frac{3}{2}\sigma_0^4\right) + 3\sigma_0.
$$

容易看到

$$
\begin{aligned}
R_{21n} \leqslant &\left| P(\sqrt{n}\sigma^{-1}(\widehat{F}_n - \overline{F}) > \log^{\frac{1}{2}} n) - (1 - \Psi_n(\log^{\frac{1}{2}} n)) \right| \\
&+ \left| P(\sqrt{n}\sigma^{-1}(\widehat{F}_n - \overline{F}) < -\log^{\frac{1}{2}} n) - \Psi_n(-\log^{\frac{1}{2}} n) \right|
\end{aligned}
$$

$$+ \left|1 - \Psi_n(\log^{\frac{1}{2}} n)\right| + \left|\Psi_n(-\log^{\frac{1}{2}} n)\right|. \quad (1.1.52)$$

根据 Chang[17] 中引理 3 及 Chow 与 Teicher[22] 中 $P_{49}$ 引理 3,可得

$$R_{21n} \leqslant C(n\log n)^{-\frac{1}{2}}. \quad (1.1.53)$$

应用不等式 $|Q(x)| \leqslant 2x^2, |x| < \frac{1}{2}$,我们有

$$R_{22n} \leqslant P\left(2\left(\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right)^2 > n^{-\frac{1}{2}}\log^{-1} n, \left|\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right| \leqslant \frac{1}{2}\right)$$

$$+ P\left(\left|\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right| > \frac{1}{2}\right) \stackrel{\mathrm{d}}{=} R_{221n} + R_{222n}. \quad (1.1.54)$$

由引理 1.1.3,我们有

$$R_{222n} \leqslant P\left(\frac{|\hat{\sigma}_{nJ}^2 - \sigma^2|}{\sigma(\hat{\sigma}_{nJ} + \sigma)} > \frac{1}{2}\right)$$

$$\leqslant P(|\hat{\sigma}_{nJ}^2 - \sigma^2| > \frac{1}{2}\sigma^2) \leqslant Cn^{1-\frac{k}{2}}, \quad (1.1.55)$$

对任何 $k > 0$ 成立.

再次使用引理 1.1.3,类似地我们有

$$R_{221n} \leqslant P\left(2\left(\frac{\hat{\sigma}_{nJ}^2 - \sigma^2}{\sigma(\hat{\sigma}_{nJ} + \sigma)}\right)^2 > n^{-\frac{1}{2}}\log^{-1} n, \hat{\sigma}_{nJ} > \frac{1}{2}\sigma\right)$$

$$\leqslant P(|\hat{\sigma}_{nJ}^2 - \sigma^2| > \frac{3}{2\sqrt{2}}\sigma^2 n^{-\frac{1}{4}}\log^{-\frac{1}{2}} n)$$

$$\leqslant Cn^{1-\frac{k}{4}}\log^{\frac{k}{2}} n, \quad (1.1.56)$$

因此,由(1.1.51),(1.1.53),(1.1.54),(1.1.56)与(1.1.55),即得

$$P(|R_{2n}| > (n\log n)^{-\frac{1}{2}}) \leqslant C((n\log n)^{-\frac{1}{2}} + n^{1-\frac{k}{4}}\log^{\frac{k}{2}} n). \quad (1.1.57)$$

由 Taylor 公式,我们有

$$R_{1n} = \sqrt{n}\sigma^{-1}\overline{F}\left[(\log\widehat{\overline{F}}_n - \log\overline{F}) + \frac{(\log\widehat{\overline{F}}_n - \log\overline{F})^2}{2}\right.$$

$$+\frac{e^{\theta'(\log\widehat{\overline{F}}_n-\log\overline{F})}}{6}(\log\widehat{\overline{F}}_n-\log\overline{F})^3\Big]\Big(1-\frac{\hat{\sigma}_{nJ}-\sigma}{\sigma}\Big)$$

$$=\sqrt{n}\sigma^{-1}\overline{F}\Big[(\log\widehat{\overline{F}}_n-\log\overline{F})\Big(1-\frac{\hat{\sigma}_{nJ}-\sigma}{\sigma}\Big)+\frac{(\log\widehat{\overline{F}}_n-\log\overline{F})^2}{2}\Big]$$

$$-\frac{\sqrt{n}\overline{F}(\hat{\sigma}_{nJ}-\sigma)(\log\widehat{\overline{F}}_n-\log\overline{F})^2}{2\sigma^2}$$

$$+\sqrt{n}\sigma^{-1}\overline{F}\,\frac{e^{\theta'(\log\widehat{\overline{F}}_n-\log\overline{F})}}{6}(\log\widehat{\overline{F}}_n-\log\overline{F})^3\Big(1-\frac{\hat{\sigma}_{nJ}-\sigma}{\sigma}\Big)$$

$$\stackrel{d}{=}R_{11n}+R_{12n}+R_{13n}. \tag{1.1.58}$$

利用 Chang[17] 中定理 2 并使用证(1.1.53) 同样的方法可证

$$P(\sqrt{n}\sigma^{-1}\overline{F}\,|(\log\widehat{\overline{F}}_n-\log\overline{F})|>\log^{\frac{1}{2}}n)\leqslant C(n\log n)^{-\frac{1}{2}}. \tag{1.1.59}$$

因此

$$P(|R_{12n}|>(n\log n)^{-\frac{1}{2}})$$

$$\leqslant P\Big(\frac{|(\hat{\sigma}_{nJ}-\sigma)(\log\widehat{\overline{F}}_n-\log\overline{F})|}{2\sigma}>n^{-\frac{1}{2}}\log^{-1}n\Big)$$

$$+P(|\sqrt{n}\sigma^{-1}\overline{F}(\log\widehat{\overline{F}}_n-\log\overline{F})|>\log^{\frac{1}{2}}n)$$

$$\leqslant P\Big(\frac{|(\hat{\sigma}_{nJ}^2-\sigma^2)(\log\widehat{\overline{F}}_n-\log\overline{F})|}{2\sigma(\hat{\sigma}_{nJ}+\sigma)}>n^{-\frac{1}{2}}\log^{-1}n,\hat{\sigma}_{nJ}>\frac{1}{2}\sigma\Big)$$

$$+P(\hat{\sigma}_{nJ}\leqslant\frac{1}{2}\sigma)+C(n\log n)^{-\frac{1}{2}}$$

$$\stackrel{d}{=}R_{121n}+R_{122n}+C(n\log n)^{-\frac{1}{2}}. \tag{1.1.60}$$

类似于证(1.1.55),可证

$$R_{122n}\leqslant Cn^{1-\frac{k}{2}}. \tag{1.1.61}$$

由(1.1.19) 与引理 1.1.3,我们有

$$R_{121n} \leqslant P(|(\hat{\sigma}_{nJ}^2 - \sigma^2)(\log\widehat{\bar{F}}_n - \log\bar{F}| > 3\sigma^2 n^{-\frac{1}{2}}\log^{-1} n)$$

$$\leqslant P(|\hat{\sigma}_{nJ}^2 - \sigma^2| > \sigma n^{-\frac{1}{4}}\log^{-\frac{1}{2}} n)$$

$$+ P(|\log\widehat{\bar{F}}_n - \log\bar{F}| > \sigma n^{-\frac{1}{4}}\log^{-\frac{1}{2}} n)$$

$$\leqslant C n^{1-\frac{k}{4}}\log^{\frac{k}{2}} n. \tag{1.1.62}$$

因此,(1.1.60),(1.1.61) 与(1.1.62) 一起就证明了

$$P(|R_{12n}| > (n\log n)^{-\frac{1}{2}}) \leqslant C(n^{1-\frac{k}{4}}\log^{\frac{k}{2}} n + (n\log n)^{-\frac{1}{2}}). \tag{1.1.63}$$

而且,我们有

$$P(|R_{13n}| > (n\log n)^{-\frac{1}{2}})$$

$$\leqslant P\left(\frac{e^{|\log\widehat{\bar{F}}_n - \log\bar{F}|}}{6}\sqrt{n}\sigma^{-1}\bar{F}\left|(\log\widehat{\bar{F}}_n - \log\bar{F})^3\right.\right.$$

$$\left.\left.\cdot\left(1 - \frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right)\right| > (n\log n)^{-\frac{1}{2}}\right)$$

$$\leqslant P\left(|\log\widehat{\bar{F}}_n - \log\bar{F}| \leqslant 1, \left|\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right| \leqslant \frac{1}{2},\right.$$

$$\left.|\log\widehat{\bar{F}}_n - \log\bar{F}|^3 > \frac{4\sigma}{e}(n\log n)^{-1}\right)$$

$$+ P(|\log\widehat{\bar{F}}_n - \log\bar{F}| > 1) + P\left(\left|\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma}\right| > \frac{1}{2}\right)$$

$$\stackrel{d}{=} R_{131n} + R_{132n} + R_{133n}. \tag{1.1.64}$$

由(1.1.19),得

$$R_{131n} \leqslant P(|\log\widehat{\bar{F}}_n - \log\bar{F}| > (n\log n)^{-\frac{1}{3}}) \leqslant C n^{-\frac{k}{6}}\log^{\frac{k}{3}} n. \tag{1.1.65}$$

类似地

$$R_{132n} \leqslant C n^{-\frac{k}{2}}. \tag{1.1.66}$$

因此,综合(1.1.64) - (1.1.66) 与(1.1.55),即得

$$P(|R_{13n}| > (n\log n)^{-\frac{1}{2}}) \leqslant C(n^{-\frac{k}{6}}\log^{\frac{k}{3}} n + n^{1-\frac{k}{2}}). \tag{1.1.67}$$

因

$$\frac{\hat{\sigma}_{nJ} - \sigma}{\sigma} = \frac{\hat{\sigma}_{nJ}^2 - \sigma^2}{2\sigma^2}\left[1 - \frac{\hat{\sigma}_{nJ} - \sigma}{2\sigma} + Q\left(\frac{\hat{\sigma}_{nJ} - \sigma}{2\sigma}\right)\right],$$

我们有

$$\begin{aligned} R_{11n} &= \sqrt{n}\sigma^{-1}\bar{F}\Big[(\log \widehat{F}_n - \log\bar{F}) \\ &\quad - (\log \widehat{F}_n - \log\bar{F})\frac{\hat{\sigma}_{nJ}^2 - \sigma^2}{2\sigma^2} + \frac{(\log \widehat{F}_n - \log\bar{F})^2}{2}\Big] \\ &\quad + \sqrt{n}\sigma^{-1}\bar{F}(\log \widehat{F}_n - \log\bar{F})\left[\frac{\hat{\sigma}_{nJ} - \sigma}{2\sigma} - Q\left(\frac{\hat{\sigma}_{nJ} - \sigma}{2\sigma}\right)\right]\frac{\hat{\sigma}_{nJ}^2 - \sigma^2}{2\sigma^2} \\ &\stackrel{\mathrm{d}}{=} R_{111n} + R_{112n}. \end{aligned} \tag{1.1.68}$$

使用证(1.1.57),(1.1.63)的方法,类似可证

$$P(|R_{112n}| > (n\log n)^{-\frac{1}{2}}) \leqslant C(n^{1-\frac{k}{4}}\log^{\frac{k}{2}} n + (n\log n)^{-\frac{1}{2}}), \tag{1.1.69}$$

而由引理 1.1.4 与引理 1.1.1,我们有

$$\begin{aligned} R_{111n} &= \sqrt{n}\sigma^{-1}\bar{F}\Big[n^{-2}\sum\sum_{i<j} h_0(Z_i,\delta_i;\ Z_j,\delta_j) \\ &\quad - \frac{1}{2}\sigma_0^2 n^{-1} + \Delta_{1n} - \Big(n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i) + r_{n1}\Big)\frac{\hat{\sigma}_{nJ}^2 - \sigma^2}{2\sigma^2} \\ &\quad + \frac{1}{2}\Big(n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i) + r_{n1}\Big)^2\Big] \\ &= \sqrt{n}\sigma_0^{-1}\Big[n^{-2}\sum\sum_{i<j} h_0(Z_i,\delta_i;\ Z_j,\delta_j) \\ &\quad - n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)\frac{\hat{\sigma}_{nJ}^2 - \sigma^2}{2\sigma^2} \\ &\quad + \frac{1}{2}n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n}\varphi(Z_i,\delta_i)\varphi(Z_j,\delta_j) - \frac{1}{2}\sigma_0 n^{-1}\Big] \end{aligned}$$

$$+\sqrt{n}\sigma_0^{-1}\Delta_{1n}-\sqrt{n}\sigma_0^{-1}r_{n1}\frac{\hat{\sigma}_{nJ}^2-\sigma^2}{2\sigma^2}$$

$$+\sqrt{n}\sigma_0^{-1}r_{n1}\frac{1}{n}\sum_{i=1}^n\varphi(Z_i,\delta_i)+\frac{\sqrt{n}\sigma_0^{-1}r_{n1}^2}{2}$$

$$\stackrel{\mathrm{d}}{=}R_{111n}^{(1)}+R_{111n}^{(2)}+R_{111n}^{(3)}+R_{111n}^{(4)}+R_{111n}^{(5)}. \quad (1.1.70)$$

由引理 1.1.4,我们有

$$P(|R_{111n}^{(2)}|>(n\log n)^{-\frac{1}{2}})=o(n^{-\frac{1}{2}}). \quad (1.1.71)$$

使用引理 1.1.1 与 1.1.3,我们得到

$$P(|R_{111n}^{(3)}|>(n\log n)^{-\frac{1}{2}})$$
$$\leqslant P(|r_{n1}|>Cn^{-\frac{3}{4}}(\log n)^{-\frac{1}{2}})+P(|\hat{\sigma}_{nJ}^2-\sigma^2|>Cn^{-\frac{1}{4}})$$
$$\leqslant C(n^{-\frac{k}{4}}\log^{\frac{k}{2}}n+n^{1-\frac{k}{4}}). \quad (1.1.72)$$

类似地,由引理 1.1.1 与(1.1.21),我们有

$$P(|R_{111n}^{(4)}|>(n\log n)^{-\frac{1}{2}})$$
$$\leqslant P(|r_{n1}|>Cn^{-\frac{3}{4}}\log^{-\frac{1}{2}}n)+P(|n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)|>Cn^{-\frac{1}{4}})$$
$$\leqslant Cn^{-\frac{k}{4}}\log^{\frac{k}{2}}n, \quad (1.1.73)$$

且

$$P(|R_{111n}^{(5)}|>(n\log n)^{-\frac{1}{2}})$$
$$\leqslant P(|r_{n1}|>Cn^{-\frac{1}{2}}\log^{-\frac{1}{4}}n)\leqslant Cn^{-\frac{k}{2}}\log^{\frac{k}{4}}n. \quad (1.1.74)$$

根据引理 1.1. 2,有

$$\Big(n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)\Big)(\hat{\sigma}_{nJ}^2-\sigma^2)$$
$$=n^{-2}\bar{F}^2\Big[\sum_{i=1}^n\Big(\varphi(Z_i,\delta_i)-\frac{1}{n}\sum_{i=1}^n\varphi(Z_i,\delta_i)\Big)^2\Big]\sum_{i=1}^n\varphi(Z_i,\delta_i)$$
$$+R_nn^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)-\sigma^2n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)$$

$$= n^{-2}\overline{F}^2\sum_{i<j}\sum[\varphi^2(Z_i,\delta_i)\varphi(Z_j,\delta_j)+\varphi(Z_i,\delta_i)\varphi^2(Z_j,\delta_j)$$

$$-(\varphi(Z_i,\delta_i)+\varphi(Z_j,\delta_j))\sigma_0^2]+n^{-2}\overline{F}^2\sum_{i=1}^n\varphi^3(Z_i,\delta_i)$$

$$-\frac{\sigma^2}{(n-1)n^2}\sum_{i<j}\sum(\varphi(Z_i,\delta_i)+\varphi(Z_j,\delta_j))$$

$$-\overline{F}^2\Big(n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)\Big)^3+R_nn^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i).\quad(1.1.75)$$

因此由(1.1.70)与(1.1.35)中分别对 $R_{111n}^{(1)}$ 与 $U_n$ 的定义,我们有

$$R_{111n}^{(1)}=\sqrt{n}\sigma_0^{-1}U_n-\frac{1}{2}n^{-\frac{1}{2}}\sigma_0^{-3}E\varphi^3(Z_1,\delta_1))$$

$$-(2\sigma_0^3n^{\frac{3}{2}})^{-1}\sum_{i=1}^n(\varphi^3(Z_i,\delta_i)-E\varphi^3(Z_i,\delta_i))$$

$$+(2\sigma_0n^{\frac{3}{2}})^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)+\sqrt{n}(2\sigma_0^3)^{-1}\Big(n^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)\Big)^3$$

$$-\sqrt{n}(2\overline{F}\sigma_0^3)^{-1}R_nn^{-1}\sum_{i=1}^n\varphi(Z_i,\delta_i)$$

$$+(2n^{\frac{3}{2}}\sigma_0)^{-1}\sum_{i=1}^n(\varphi_0^2(Z_i,\delta_i)-E\varphi_0^2(z_i,\delta_i))$$

$$\stackrel{\mathrm{d}}{=}\sqrt{n}\sigma_0^{-1}U_n+e_{1n}+e_{2n}+e_{3n}+e_{4n}+e_{5n}+e_{6n}.\quad(1.1.76)$$

而使用 Tchebychev 不等式,我们得到

$$P(|e_{2n}|>C(n\log n)^{-\frac{1}{2}})$$

$$\leqslant Cn^{-2}\log nE\Big[\sum_{i=1}^n(\varphi^3(Z_i,\delta_i)-E\varphi^3(Z_i,\delta_i))\Big]^2$$

$$\leqslant Cn^{-1}\log n.\quad(1.1.77)$$

类似地

$$P(|e_{3n}|>C(n\log n)^{-\frac{1}{2}})\leqslant Cn^{-1}\log n,\quad(1.1.78)$$

$$P(|e_{6n}|>C(n\log n)^{-\frac{1}{2}})\leqslant Cn^{-1}\log n.\quad(1.1.79)$$

由(1.1.71),我们得

$$P(|e_{4n}| > C(n\log n)^{-\frac{1}{2}})$$
$$\leqslant P\left(\left|n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)\right| > Cn^{-\frac{1}{3}}(\log n)^{-\frac{1}{6}}\right)$$
$$\leqslant Cn^{-\frac{k}{6}}\log^{\frac{k}{6}} n. \tag{1.1.80}$$

容易看见

$$P(|e_{5n}| > C(n\log n)^{-\frac{1}{2}})$$
$$\leqslant P\left(\left|n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)\right| > Cn^{-\frac{2}{3}}(\log n)^{-\frac{1}{2}}\right)$$
$$+ P(|R_n| > Cn^{-\frac{1}{3}}). \tag{1.1.81}$$

注意到 $\varphi(Z_i,\delta_i),1 \leqslant i \leqslant n$ 是有界独立同分布随机变量且 $E\varphi(Z_1,\delta_1) = 0$.因此,由 Rao[64] 中命题 1.14.16 知,对任何 $k > 0$,有

$$P\left(\left|n^{-1}\sum_{i=1}^{n}\varphi(Z_i,\delta_i)\right| > Cn^{-\frac{2}{3}}(\log n)^{-\frac{1}{2}}\right)$$
$$\leqslant Cn^{-k}[n^{-\frac{2}{3}}(\log n)^{-\frac{1}{2}}]^{-1-k} \leqslant Cn^{\frac{2-k}{3}}(\log n)^{\frac{1+k}{2}}. \tag{1.1.82}$$

于是(1.1.81),(1.1.82) 与(1.1.25) 一起就证明了

$$P(|e_{5n}| > C(n\log n)^{-\frac{1}{2}} \leqslant C(n^{\frac{2-k}{3}}(\log n)^{\frac{1+k}{2}} + n^{1-\frac{k}{6}}). \tag{1.1.83}$$

综合(1.1.42),(1.1.39),(1.1.57),(1.1.58),(1.1.63),(1.1.67) - (1.1.74),(1.1.76) - (1.1.80) 与(1.1.83) 且注意到 $k$ 是任意常数,我们有

$$\sqrt{n}\hat{\sigma}_{nJ}^{-1}(\widehat{\overline{F}}_n - \overline{F}) = \sqrt{n}\sigma_0^{-1}U_n - An^{-\frac{1}{2}} + E_n,$$

且

$$P(|E_n| > C(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}),$$

其中 $A$ 如引理 1.1.8 所定义.由引理 1.1.8 与 1.1.5,定理 1.1.1 得

证.

定理 1.1.2 的证明:注意应用分部积分,我们有

$$
\begin{aligned}
&|\sigma_{0n}^2-\sigma_0^2|\\
&=\left|\int_0^t \overline{H}_n^{-2}\mathrm{d}\widetilde{H}_{n1}-\int_0^t \overline{H}^{-2}\mathrm{d}\widetilde{H}_1\right|\\
&=\left|\int_0^t (\overline{H}_n^{-2}-\overline{H}^{-2})\mathrm{d}\widetilde{H}_{n1}\right|+\left|\int_0^t \overline{H}^{-2}\mathrm{d}(\widetilde{H}_{n1}-\widetilde{H}_1)\right|\\
&\leqslant 2\overline{H}_n^{-2}(t)\overline{H}^{-2}(t)\sup_{0\leqslant s\leqslant t}|\overline{H}_n(s)-\overline{H}(s)|\\
&\quad+2\overline{H}^{-2}(t)\sup_{0\leqslant s\leqslant t}|\widetilde{H}_{n1}(s)-\widetilde{H}_1(s)|\\
&\quad+\overline{H}^{-3}(t)\sup_{0\leqslant s\leqslant t}|\widetilde{H}_{n1}(s)-\widetilde{H}(s)|\xrightarrow{a.s.}0.
\end{aligned}
\tag{1.1.84}
$$

类似可证

$$\kappa_{3n}\xrightarrow{a.s.}\kappa_3.$$

由此可得

$$\sup_x|K_n^*(x)-K_n(x)|=o(n^{-\frac{1}{2}}),\quad a.s. \tag{1.1.85}$$

于是由定理 1.1.1 与(1.1.85),即证定理 1.1.2.

## §1.2 基于左截断右删失数据乘积限估计的一些逼近定理

### 1.2.1 记号与主要结果

设 $X$ 是表示寿命或类似寿命的随机变量,且有分布函数 $F$,设 $T$ 与 $Y$ 是表示左截断与右删失随机变量且分别具有分布函数$G$ 与 $L$. 本节假定 $F,G$ 与 $L$ 是连续分布函数且$X$ 独立于$(T,Y)$,但 $T$ 与 $Y$ 可能相依. 在随机左截断右删失模型中,我们仅能在 $T\leqslant Z$ 时观察到$(Z,T,\delta)$,其中 $Z=\min(X,Y)$,$\delta=I[X\leqslant Y]$,$I[\cdot]$ 表示某事件的示性函数.设 $\nu=P(T\leqslant Z)>0$,$W$ 定义为$Z$ 的分布函数,显然,$1-W=(1-F)(1-L)$. 设$(Z_i,T_i,\delta_i)$,$i=1,2,$

$\cdots,n$ 是 $n$ 个独立同分布来自$(Z,T,\delta)$的样本观察(即 $T_i \leqslant Z_i, i = 1,2,\cdots,n$).

设

$$C_n(z) = n^{-1}\sum_{i=1}^{n} I[T_i \leqslant z \leqslant Z_i],$$

基于上面左截断右删失观察,Tsai,Jewell 与 Wang[79]定义了 $F$ 的乘积限估计(PLE)$\hat{F}_n$如下:

$$1 - \hat{F}_n(x) = \prod_{Z_i \leqslant x}(1 - [nC_n(Z_i)]^{-1})^{\delta_i}. \qquad (1.2.1)$$

正如 Gijbels 与 Wang[37] 所指出,当左截断不发生时 $\hat{F}_n$ 就变成如§1.1 所述的 KM 估计,当没有右删失(即 $Y = \infty$)时,$\hat{F}_n$ 就是 Lynden-Bell[53] 的 PL 估计.关于 KM 估计的研究已相当多(见引言中的叙述),而对 Lynden-Bell PL 估计的一些研究结果可见 Woodroof[87],Wang,Jewell 与 Tsai[83] 与 Chao 与 Lo[19].对任何分布函数 $K$,定义 $a_K = \inf\{t: K(t) > 0)$ 与 $b_K = \inf\{t: K(t) = 1\}$,且为简便计,我们约定 $\alpha$ 可表示任何所需的绝对常数,$\alpha_z$ 可表示任何可能与$z$有关的常数,既使在同一式中出现 $\alpha$(或 $\alpha_z$)均可表示不同.

本节假定 $0 \leqslant a_G \leqslant a_W, b_G \leqslant b_W \leqslant \infty$ ,且记

$$W_1(y) = P(Z \leqslant y, \delta = 1 \mid T \leqslant Z),$$

$$W_{1n}(y) = n^{-1}\sum_{i=1}^{n} I[Z_i \leqslant y, \delta_i = 1],$$

$$C(z) = P(T \leqslant z \leqslant Z \mid T \leqslant Z),$$

$$\beta_k(z) = \int_0^z \frac{\mathrm{d}W_1(u)}{C^k(u)}, k = 1.2,\cdots$$

$$\sigma_1^2 = \beta_2(z), \quad \sigma^2 = (1-F)^2\sigma_1^2,$$

$$\Lambda(z) = \int_0^z \frac{\mathrm{d}F(u)}{1 - F(u)},$$

其中 $\Lambda(z)$ 定义为 $F$ 的累积失效率函数.由 Gijbels 与 Wang[37],

$\Lambda(z)$ 能表示成

$$\Lambda(z) = \int_0^z \frac{\mathrm{d}W_1(u)}{C(u)},$$

因此 $\Lambda(z)$ 的一个自然估计为

$$\Lambda_n(z) = \int_0^z \frac{\mathrm{d}W_{1n}(u)}{C_n(u)}.$$

其中 $C_n(z)$ 如前所定义，是 $C(z)$ 的经验估计. 设 $U^{(n)} = (1 - F(z))U_n$，与

$$U_n = \frac{1}{n^2}\sum_{i<j}^{n} h((Z_i, T_i, \delta_i), (Z_j, T_j, \delta_j)),$$

对 $i, j = 1, 2, \cdots, n$,

$$\begin{aligned}
&h((Z_i, T_i, \delta_i), (Z_j, T_j, \delta_j)) \\
&= \zeta(Z_i, \delta_i, T_i) + \zeta(Z_j, \delta_j, T_j) \\
&\quad - \frac{\zeta(Z_i, \delta_i, T_i)\zeta(Z_j, \delta_j, T_j) + \zeta(Z_j, \delta_j, T_j)\zeta(Z_i, \delta_i, T_i)}{2} \\
&\quad + \psi((Z_i, T_i, \delta_i), (Z_j, T_j, \delta_j)) + \psi((Z_j, T_j, \delta_j), (Z_i, T_i, \delta_i)), \\
&\zeta(Z_i, \delta_i, T_i) = \zeta_1(Z_i)\delta_i + \zeta_2(T_i, Z_i), \\
&\zeta_1(Z_i) = C^{-1}(Z_i) I[Z_i \leqslant z], \\
&\zeta_2(T_i, Z_i) = -\int_0^z C^{-2}(u) I[T_i \leqslant u \leqslant Z_i] \mathrm{d}W_1(u), \\
&\psi((Z_i, T_i, \delta_i), (Z_j, T_j, \delta_j)) \\
&= \psi_1((Z_i, Z_j), (T_i, T_j))\delta_i + \psi_2((Z_i, Z_j), (T_i, T_j)) \\
&\quad - E[\psi_1((Z_i, Z_j), (T_i, T_j))\delta_i + \psi_2((Z_i, Z_j), \\
&\quad (T_i, T_j)) \mid T_i, Z_i, \delta_i], \psi_1((Z_i, Z_j), (T_i, T_j)) \\
&= -C^{-2}(Z_i) I[Z_i \leqslant z] I[T_j \leqslant Z_i \leqslant Z_j], \psi_2((Z_i, Z_j), (T_i, T_j)) \\
&= \int_0^z C^{-3}(u) I[T_i \leqslant u \leqslant Z_i] I[T_j \leqslant u \leqslant Z_j] \mathrm{d}W_1(u),
\end{aligned}$$

且定义 $Q(x) = \dfrac{1}{1+x} - (1 - x)$，于是有下面的

**定理 1.2.1** 设 $G,F$ 与 $L$ 是连续分布函数且 $a_G < a_W$,则对任何 $a_W < z < b_W$,有

$$\hat{F}_n(z) - F(z) = U^{(n)} + \sum_{i=1}^{9} R_{in},$$

其中

$$R_{1n} = (1 - F(z))\frac{1}{n^2}\sum_{i=1}^{n}\zeta(Z_i,\delta_i,T_i),$$

$$R_{2n} = (1 - F(z))\frac{1}{n^2}\sum_{i=1}^{n}\psi((Z_i,T_i,\delta_i),(Z_i,T_i,\delta_i)),$$

$$R_{3n} = (1 - F(z))\int_0^z \frac{(C_n(u) - C(u))^2}{C^3(u)}\mathrm{d}(W_{1n}(u) - W_1(u)),$$

$$R_{4n} = (1 - F(z))\int_0^z \frac{C(u) - C_n(u)}{C^2(u)}Q(\frac{C_n(u) - C(u)}{C(u)})\mathrm{d}W_{1n}(u),$$

$$R_{5n} = -(1 - F(z))(\log(1 - \hat{F}_n(z)) + \Lambda_n(z)),$$

$$R_{6n} = -(1 - F(z))\frac{1}{2n^2}\sum_{i=1}^{n}\zeta^2(Z_i,\delta_i,T_i),$$

$$R_{7n} = -(1 - F(z))\frac{1}{n}[\sum_{i=1}^{n}\zeta(Z_i,\delta_i,T_i)]r_n,$$

$$R_{8n} = -(1 - F(z))\frac{1}{2}r_n^2,$$

$$R_{9n} = -(1 - F(z))\frac{(\log(1 - \hat{F}_n(z)) + \Lambda(z))^3}{6} \cdot \mathrm{e}^{\theta(\log(1-\hat{F}_n(z))+\Lambda(z))}, \quad 0 < \theta < 1,$$

且 $R_{7n},R_{8n}$ 中的 $r_n$ 如下面引理 1.2.1 中所定义.

由上面表示定理及 $U$ 统计量的 Berry-Essen 不等式,通过证明 $nR_{in}(i = 1,2,\cdots,9)$ 的尾概率不超过 $\alpha_z n^{-\frac{1}{2}}$ 可证下面定理,其中 $\alpha_z$ 是待定的可能与 $z$ 有关的常数.

**定理 1.2.2** 设 $G,F$ 与 $L$ 是连续的分布函数且 $a_G < a_W$,则对任何取定的 $a_W < z < b_W$,有

$$\sup_x |P(\sqrt{n}\sigma^{-1}(\hat{F}_n(z) - F(z)) \leqslant x) - \Phi(x)|$$
$$\leqslant \alpha_0 \sigma_1^{-3} \beta_{14}(z) n^{-\frac{1}{2}},$$

其中 $\Phi$ 是标准正态分布函数，$\sigma, \sigma_1, \beta_{14}(z)$ 如前所定义.

由 Gijbels 与 Wang[37] 的定理 1(c) 可证

**定理 1.2.3** 设 $G, F$ 与 $L$ 是连续分布函数，$a_G < a_W$，则对任何 $a_W < z < b_W$ 及 $r > 1$，有

$$E|\hat{F}_n(z) - F(z)|^r$$
$$\leqslant 2^{2r-1}(1 - F(z))^r C^{-r}(a_W) n^{-\frac{r}{2}} + O(n^{-r}),$$

**定理 1.2.4** 设 $G, F$ 与 $L$ 是连续分布函数，$a_G < a_W$，则对任何 $a_W < z < b_W$，有

$$\overline{\lim_n} \frac{\sqrt{n}(\hat{F}_n(z) - F(z))}{(\log\log n)^{1/2}} = \sqrt{2}\sigma, \quad a.s.$$

代替使用 Gijbels 与 Wang[37] 中的定理 1(c)，本节定理 1.2.1 也能用于证明定理 1.2.3 与定理 1.2.4.

### 1.2.2 定理的证明

**引理 1.2.1** 若 $G, F$ 与 $L$ 是连续分布函数且 $a_G < a_W$，则对任何 $a_W < z < b_W$，有

(i) $$-(\log(1 - \hat{F}_n(z)) + \Lambda(z)) = \frac{1}{n}\sum_{i=1}^{n} \zeta(Z_i, \delta_i, T_i) + r_n,$$

且

$$P(n|r_n| > x)$$
$$\leqslant 2ne^{-\alpha n} + K[e^{-\alpha x} + (x/5)^{-2n} + e^{-\lambda(nx)^{1/2}} + e^{-\lambda x} + e^{-\lambda n x^2}],$$

其中 $K, \lambda$ 均表示绝对常数.

(ii) $$\hat{F}_n(z) - F(z) = (1 - F(z))\frac{1}{n}\sum_{i=1}^{n} \zeta(Z_i, \delta_i, T_i) + R_n,$$ 且

$$R_n(z) = O(n^{-1}\log n), \quad a.s.,$$
$$E|R_n(z)|^r = O(n^{-r}),$$

对任何 $r > 0$ 成立.

证明:注意到

$$-(\log(1-\hat{F}_n(z)) + \Lambda(z))$$
$$= \Lambda_n(z) - \Lambda(z) - (\log(1-\hat{F}_n(z)) + \Lambda_n(z)).$$

综合上式与 Gijbels 与 Wang[37] 中的(3.9),(3.10),(3.2),(1.7),(1.8) 及其中的引理 2 - 引理 4,引理 1.2.1(i) 即得证. 而引理 1.2.1(ii) 正是 Gijbels 与 Wang[37] 中的定理 1(c). 于是完成了引理 1.2.1 的证明.

定理 1.2.1 的证明: 应用三项 Taylor 公式及引理 1.2.1(i),我们有

$$\hat{F}_n(z) - F(z) = -(e^{\log(1-\hat{F}_n(z))} - e^{\log(1-F(z))})$$
$$= -(1-F(z))[\log(1-\hat{F}_n(z)) + \Lambda(z)]$$
$$-(1-F(z))\frac{(\log(1-\hat{F}_n(z)+\Lambda(z))^2}{2}$$
$$-(1-F(z))\frac{(\log(1-\hat{F}_n(z))+\Lambda(z))^3}{6}$$
$$\cdot e^{\theta(\log(1-\hat{F}_n(z))+\Lambda(z))} \qquad 0<\theta<1$$
$$= (1-F(z))(\Lambda_n(z)-\Lambda(z))$$
$$-(1-F(z))(\log(1-\hat{F}_n(z))+\Lambda_n(z)) - (1-F(z))$$
$$\frac{1}{n^2}\sum_{i<j}\frac{\zeta(Z_i,\delta_i,T_i)\zeta(Z_j,\delta_j,T_j)+\zeta(Z_j,\delta_j,T_j)\zeta(Z_i,\delta_i,T_i)}{2}$$
$$-(1-F(z))\frac{1}{2n^2}\sum_{i=1}^{n}\zeta^2(Z_i,\delta_i,T_i)$$
$$-(1-F(z))\frac{r_n}{n}\sum_{i=1}^{n}\zeta(Z_i,\delta_i,T_i) - \frac{(1-F(z)}{2}r_n^2$$
$$-\frac{(1-F(z))}{6}(\log(1-\hat{F}_n(z))+\Lambda(z))^3 \times e^{\theta(\log(1-\hat{F}_n(z)+\Lambda(z))}$$

$$= (1 - F(z))(\Lambda_n(z) - \Lambda(z)) - (1 - F(z))$$
$$\frac{1}{n^2}\sum_{i<j}\frac{\zeta(Z_i,\delta_i,T_i)\zeta(Z_j,\delta_j,T_j) + \zeta(Z_j,\delta_j,T_j)\zeta(Z_i,\delta_i,T_i)}{2}$$
$$+ R_{5n} + R_{6n} + R_{7n} + R_{8n} + R_{9n}, \tag{1.2.2}$$

其中 $R_{5n},\cdots,R_{9n}$ 如定理 1.2.1 所定义.

下面我们考虑

$$\Lambda_n(z) - \Lambda(z)$$
$$= \sum_{i=1}^{n}\frac{I[Z_i \leqslant z, \delta_i = 1]}{nC_n(Z_i)} - \int_0^z \frac{\mathrm{d}W_1(u)}{C(u)}$$
$$= I_n + \sum_{i=1}^{n}\frac{I[Z_i \leqslant z, \delta_i = 1]}{nC(Z_i)} - \int_0^z \frac{\mathrm{d}W_1(u)}{C(u)}, \tag{1.2.3}$$

其中

$$I_n = \frac{1}{n}\sum_{i=1}^{n}\frac{I[Z_i \leqslant z, \delta_i = 1](C(Z_i) - C_n(Z_i))}{C_n(Z_i)C(Z_i)}.$$

设 $Q(x)$ 如节 1.2.1 中所定义,则

$$I_n = \frac{1}{n}\sum_{i=1}^{n}\frac{I[Z_i \leqslant z, \delta_i = 1](C(Z_i) - C_n(Z_i))}{C^2(Z_i)}$$
$$+ \frac{1}{n}\sum_{i=1}^{n}\frac{I[Z_i \leqslant z, \delta_i = 1](C(Z_i) - C_n(Z_i))^2}{C^3(Z_i)}$$
$$+ \frac{1}{n}\sum_{i=1}^{n}\frac{I[Z_i \leqslant z, \delta_i = 1](C(Z_i) - C_n(Z_i))}{C^2(Z_i)}Q\left(\frac{C_n(Z_i) - C(Z_i)}{C(Z_i)}\right)$$
$$\stackrel{\mathrm{d}}{=} I_{1n} + I_{2n} + I_{3n}. \tag{1.2.4}$$

易见

$$I_{1n} = -\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left\{\frac{I[Z_i \leqslant z, \delta_i = 1]I[T_j \leqslant Z_i \leqslant Z_j]}{C^2(Z_i)}\right.$$
$$\left. - E\left[\frac{I[Z_i \leqslant z, \delta_i = 1]I[T_j \leqslant Z_i \leqslant Z_j]}{C^2(Z_i)}\,\middle|\, T_i, Z_i, \delta_i\right]\right\}. \tag{1.2.5}$$

(注: 在左截断右删失模型中随机观察函数的期望是关于 $T \leqslant Z$

的条件下联合条件分布函数而求的,即对任何使得 $E|\varphi(T,Z))| < \infty$ 的 $\varphi(\cdot,\cdot)$,有 $E\varphi(T,Z) = \iint \varphi(t,y)\mathrm{d}P(T \leqslant t, Z \leqslant y | T \leqslant Z)$.)

容易验证

$$I_{2n} = \int_0^z \frac{(C_n(u) - C(u))^2}{C^3(u)} \mathrm{d}W_1(u) + \int_0^z \frac{(C_n(u) - C(u))^2}{C^3(u)} \mathrm{d}(W_{1n}(u) - W_1(u)) \stackrel{\mathrm{d}}{=} I_{21n} + \frac{R_{3n}}{1 - F(z)}, \tag{1.2.6}$$

其中 $R_{3n}$ 如定理 1.2.1 所定义.注意到

$$\frac{1}{n}\sum_{i=1}^n \int_0^z \frac{I[T_i \leqslant u \leqslant Z_i]}{C^2(u)} \mathrm{d}W_1(u) = \frac{1}{n}\sum_{i=1}^n E\left[\int_0^z \frac{I[T_i \leqslant u \leqslant Z_i] I[T_j \leqslant u \leqslant Z_j]}{C^3(u)} \cdot \mathrm{d}W_1(u) \,\Big|\, T_i, Z_i, \delta_i\right], \tag{1.2.7}$$

因此

$$\begin{aligned} I_{21n} &= \int_0^z \frac{\left(\frac{1}{n}\sum_{j=1}^n I[T_j \leqslant u \leqslant Z_j] - C(u)\right)^2}{C^3(u)} \mathrm{d}W_1(u) \\ &= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \left\{\int_0^z \frac{I[T_i \leqslant u \leqslant Z_i] I[T_j \leqslant u \leqslant Z_j]}{C^3(u)} \mathrm{d}W_1(u) \right. \\ &\quad \left. - E\left[\int_0^z \frac{I[T_i \leqslant u \leqslant Z_i] I[T_j \leqslant u \leqslant Z_j]}{C^3(u)} \mathrm{d}W_1(u) \,\Big|\, T_i, Z_i, \delta_i\right]\right\} \\ &\quad - \frac{1}{n}\sum_{i=1}^n \int_0^z \frac{I[T_i \leqslant u \leqslant Z_i]}{C^2(u)} \mathrm{d}W_1(u) + \int_0^z \frac{\mathrm{d}W_1(u)}{C(u)}. \end{aligned} \tag{1.2.8}$$

易见

$$I_{3n} = \frac{R_{4n}}{1 - F(z)}, \tag{1.2.9}$$

其中 $R_{4n}$ 如定理 1.2.1 所定义.

综合(1.2.3) - (1.2.6),(1.2.8) 及(1.2.9),可得

$$\begin{aligned}\Lambda_n(z)-\Lambda(z)&=\frac{1}{n}\sum_{i=1}^{n}\frac{I[Z_i\leqslant z,\delta_i=1]}{C(Z_i)}\\&-\frac{1}{n}\sum_{i=1}^{n}\int_0^z\frac{I[T_i\leqslant u\leqslant Z_i]}{C^2(u)}\mathrm{d}W_1(u)\\&+\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\bigg\{-\frac{I[Z_i\leqslant z,\delta_i=1]I[T_j\leqslant Z_i\leqslant Z_j]}{C^2(Z_i)}\\&+\int_0^z\frac{I[T_i\leqslant u\leqslant Z_i]I[T_j\leqslant u\leqslant Z_j]}{C^3(u)}\mathrm{d}W_1(u)\\&-E\left[-\frac{I[Z_i\leqslant z,\delta_i=1]I[T_j\leqslant Z_i\leqslant Z_j]}{C^2(Z_i)}\,\middle|\,T_i,Z_i,\delta_i\right]\\&-E\left[\int_0^z\frac{I[T_i\leqslant u\leqslant Z_i]I[T_j\leqslant u\leqslant Z_j]}{C^3(u)}\mathrm{d}W_1(u)\,\middle|\,T_i,Z_i,\delta_i\right]\bigg\}\\&+\frac{R_{3n}}{1-F(z)}+\frac{R_{4n}}{1-F(z)}.\end{aligned}\tag{1.2.10}$$

易见

$$\begin{aligned}&\Lambda_n(z)-\Lambda(z)\\&=\frac{1}{n^2}\sum_{i<j}[\zeta(Z_i,\delta_i,T_i)+\zeta(Z_j,\delta_j,T_j)\\&\quad+\psi((Z_i,T_i,\delta_i),(Z_j,T_j,\delta_j))+\psi((Z_j,T_j,\delta_j),(Z_i,T_i,\delta_i))]\\&\quad+\frac{R_{1n}}{1-F(z)}+\frac{R_{2n}}{1-F(z)}+\frac{R_{3n}}{1-F(z)}+\frac{R_{4n}}{1-F(z)},\end{aligned}\tag{1.2.11}$$

其中 $R_{1n},R_{2n}$ 也如定理 1.2.1 所定义,易见(1.2.2) 与(1.2.11)一起证明了定理 1.2.1.

**引理 1.2.2** 对任何随机变量 $\xi,\eta$ 与实常数 $a>0$,有

$$\begin{aligned}&\sup_x|P(\xi+\eta\leqslant x)-\Phi(x)|\\&\leqslant\sup_x|P(\xi\leqslant x)-\Phi(x)|+\frac{a}{\sqrt{2\pi}}+P(|\xi|>a).\end{aligned}$$

**引理 1.2.3** 在定理 1.2.2 的条件下，存在一个正的绝对常数 $\alpha$，使得对任何取定的 $a_W < z < b_W$，有

$$\sup_x |P(\sqrt{n}\sigma^{-1}U^{(n)} \leqslant x) - \Phi(x)| \leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}},$$

其中 $\beta_{14}(z)$ 如 1.2.1 节所定义.

证明：注意到

$$\sqrt{n}\sigma^{-1}U^{(n)} = \sqrt{n}\sigma_1^{-1}U_n,$$

因此，我们仅需证明 $U_n$ 有引理 1.2.3 的结果.

既然

$$E\zeta(Z_1,\delta_1,T_1) = 0,\quad E\psi((Z_1,T_1,\delta_1),(Z_2,T_2,\delta_2)) = 0, \tag{1.2.12}$$

$U_n$ 是为 2 阶且具有对称核的 $U$ 统计量，且 $Eh = 0$.

直接计算得到

$$E\zeta^2(Z_1,\delta_1,T_1) = \int_0^z \frac{\mathrm{d}W_1(u)}{C^2(u)} = \sigma_1^2. \tag{1.2.13}$$

由 $C_3$ 不等式，我们有

$$\begin{aligned} &E|\psi((Z_1,T_1,\delta_1),(Z_2,T_2,\delta_2))|^3 \\ &\leqslant 4\{E|\psi_1((Z_1,Z_2),(T_1,T_2))\delta_1|^3 \\ &\quad + E|\psi_2((Z_1,Z_2),(T_1,T_2))|^3\}, \end{aligned} \tag{1.2.14}$$

其中

$$\begin{aligned} &E|\psi_1((Z_1,Z_2),(T_1,T_2))\delta_1|^3 \\ &= E\left[\frac{I[Z_1 \leqslant z]I[T_2 \leqslant Z_1 \leqslant Z_2]\delta_1}{C^6(Z_1)}\,\bigg|\, Z_1,\delta_1\right] \\ &= E\left[\frac{I[Z_1 \leqslant z,\delta_1 = 1]}{C^5(Z_1)}\right] = \beta_5(z), \end{aligned} \tag{1.2.15}$$

且

$$\begin{aligned} &E|\psi_2((Z_1,Z_2),(T_1,T_2))|^3 \\ &= E\left[\int_0^z \frac{I[T_1 \leqslant u \leqslant Z_1]I[T_2 \leqslant u \leqslant Z_2]}{C^3(u)}\mathrm{d}W_1(u)\right]^3 \end{aligned}$$

$$\leqslant \int_0^z \frac{E([T_1 \leqslant u \leqslant Z_1]I[T_2 \leqslant u \leqslant Z_2])}{C^9(u)} \mathrm{d}W_1(u)$$

$$= \beta_7(z). \tag{1.2.16}$$

由(1.2.14) - (1.2.16),得

$$E|\psi((Z_1, T_1, \delta_1), (Z_2, T_2, \delta_2))|^3 \leqslant 8\beta_7(z). \tag{1.2.17}$$

直接计算,得

$$\begin{aligned} &\mathrm{E}|\zeta(Z_1, \delta_1, T_1)|^3 \\ &= E\{\zeta_1^3(Z_1)\delta_1\} + 3E\{\zeta_1^2(Z_1)|\zeta_2(T_1, Z_1)|\delta_1\} \\ &\quad + 3E\zeta_1(Z_1)|\zeta_2^2(T_1, Z_1)| + E|\zeta_2^3(T_1, Z_1)| \\ &\leqslant \beta_3(z) + 3\beta_2^2(z) + 3\beta_1(z)\beta_2^2 + \beta_2^3 \leqslant 8\beta_7(z). \end{aligned} \tag{1.2.18}$$

因此,(1.2.17) 与(1.2.18) 一起证明了

$$E|h((Z_1, T_1, \delta_1), (Z_2, T_2, \delta_2))|^3 \leqslant 465\beta_{14}(z). \tag{1.2.19}$$

易见

$$\begin{aligned} h_1(Z_1, T_1, \delta_1) &\stackrel{\mathrm{d}}{=} E[h((Z_1, T_1, \delta_1), (Z_2, T_2, \delta_2) | Z_1, T_1, \delta_1)] \\ &= \zeta(Z_1, T_1, \delta_1), \end{aligned}$$

因此

$$\tau \stackrel{\mathrm{d}}{=} \mathrm{Var} h_1(Z_1, T_1, \delta_1) = \sigma_1^2. \tag{1.2.20}$$

由 Callaert 与 Janssen[13],存在正的绝对常数 $\alpha$,使得

$$\sup_x \left| P\left( \frac{\frac{2n}{n-1}\sqrt{n}U_n}{2\tau^{1/2}} \leqslant x \right) - \Phi(x) \right| \leqslant \alpha E|h|^3 (4\tau)^{-\frac{3}{2}} n^{-\frac{1}{2}},$$

由此及(1.2.19) 与(1.2.20),得

$$\sup_x \left| P\left( \frac{n}{n-1}\sqrt{n}\sigma_1^{-1} U_n \leqslant x \right) - \Phi(x) \right| \leqslant \alpha\sigma_1^{-3}\beta_{14}(z) n^{-\frac{1}{2}}. \tag{1.2.21}$$

注意到

$$\frac{n}{n-1}\sqrt{n}U_n = \sqrt{n}U_n + \frac{1}{n-1}\sqrt{n}U_n. \tag{1.2.22}$$

由 Tchebychev 不等式和 $U$ 统计量的方差公式(见 Serfling[70]),我们有

$$P\left(\frac{1}{n-1}\sqrt{n}U_n > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}\right) \leqslant \alpha_z \frac{n^2}{(n-1)^2}EU_n^2$$

$$\leqslant \alpha_z\left[\frac{\binom{2}{1}\binom{n-2}{1}}{\binom{n}{2}}\tau + \frac{\binom{2}{2}\binom{n-2}{1}}{\binom{n}{2}}Eh^2((Z_1,T_1,\delta_1),(Z_2,T_2,\delta_2))\right]$$

$$\leqslant \alpha_z n^{-1} \leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}. \tag{1.2.23}$$

根据引理 1.2.2 与(1.2.21) - (1.2.23),得

$$|P(\sqrt{n}\sigma_1^{-1}U_n \leqslant x) - \Phi(x)| \leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}, \tag{1.2.24}$$

这就证明了引理 1.2.3.

为使用定理 1.2.1,引理 1.2.2、1.2.3 证明定理 1.2.2,剩下需证明

**引理 1.2.4** 在定理 1.2.2 的条件下,对任何 $a_W < z < b_W$,我们有

$$P(\sqrt{n}\sigma^{-1}|R_{in}| > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}) \leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}},$$

其中 $i = 1,2,\cdots,9$.

我们将在 1.2.3 中证明引理 1.2.4.

定理 1.2.2 的证明:设

$$\widetilde{R}_n(z) = \sum_{i=1}^{9} R_{in}. \tag{1.2.25}$$

由引理 1.2.4,得

$$P(\sqrt{n}\sigma^{-1}|\widetilde{R}_n(z)| > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}})$$

$$\leqslant \sum_{i=1}^{9} P(\sqrt{n}\sigma^{-1}|R_{in}| > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}})$$

$$\leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}. \tag{1.2.26}$$

综合定理 1.2.1,(1.2.25),(1.2.26) 且利用引理 1.2.2 与 1.2.3,定理 1.2.2 即得证.

定理 1.2.3 的证明：由引理 1.2.1(ii),Tchebychev 不等式与 Dharmadhikari-Jogdeo(D-J) 不等式(见 Rao[64]),对任何 $r>1$,有

$$
\begin{aligned}
& E|\hat{F}_n(z)-F(z)|^r \\
\leqslant\ & 2^{r-1}\{(1-F(z))^r E[\frac{1}{n}\sum_{i=1}^{n}\zeta(Z_i,\delta_i,T_i)]^r+E|R_n|^r\} \\
\leqslant\ & 2^{r-1}(1-F(z))^r n^{-r} n^{\frac{r}{2}} E|\zeta(Z_1,\delta_1,T_1)|^r+O(n^{-r}),
\end{aligned}
\tag{1.2.27}
$$

且

$$
\begin{aligned}
E|\zeta(Z_1,\delta_1,T_1)|^r &\leqslant 2^{r-1}\left(\int_0^z \frac{\mathrm{d}W(u)}{C^r(u)}+\int_0^z \frac{\mathrm{d}W_1(u)}{C^{r-1}(u)}\right) \\
&\leqslant 2^r C^{-r}(a_W)<\infty.
\end{aligned}
\tag{1.2.28}
$$

综合(1.2.27) 与(1.2.28),得

$$
\begin{aligned}
& E|\hat{F}_n(z)-F(z)|^r \\
\leqslant\ & 2^{2r-1}(1-F(z))^r C^{-r}(a_W) n^{-\frac{r}{2}}+O(n^{-r}),
\end{aligned}
$$

于是定理 1.2.3 得证.

定理 1.2.4 的证明：既然 $E\zeta=0, E\zeta^2=\sigma_1^2\leqslant C^{-2}(a_W)<\infty$,于是由独立同分布随机变量和的重对数律及引理 1.2.1(ii),我们得

$$
\begin{aligned}
& \overline{\lim_n}\frac{\sqrt{n}(\hat{F}_n(z)-F(z))}{\sqrt{\log\log n}} \\
=\ & (1-F(z))\overline{\lim_n}\frac{\sqrt{n}(n^{-1}\sum_{i=1}^{n}\zeta(Z_i,\delta_i,T_i))}{\sqrt{\log\log n}} \\
& +\overline{\lim_n}\frac{\sqrt{n}R_n}{\sqrt{\log\log n}}\rightarrow\sqrt{2}\sigma.
\end{aligned}
$$

这就证明了定理 1.2.4.

### 1.2.3 引理 1.2.4 的证明

由 Tchebychev 不等式,我们有

$$P(\sqrt{n}\sigma^{-1}|R_{1n}| > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}})$$

$$\leqslant \alpha n^2 n^{-4} nE\zeta^2(Z_1,\delta_1,T_1) \leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}, \quad (1.2.29)$$

这就证明了 $i=1$ 的情形.

注意到

$$\begin{aligned}
&E|\psi((Z_i,T_i,\delta_i),(Z_i,T_i,\delta_i))|\\
\leqslant &|E\psi_1((Z_1,Z_1),(T_1,T_1))\delta_1| + |E\psi_2((Z_1,Z_1),(T_1,T_1)|\\
&+|E\psi_1((Z_1,Z_2),(T_1,T_2))\delta_1| + |E\psi_2((Z_1,Z_2),(T_1,T_2))|\\
\leqslant &|EC^{-2}(Z_1)I[Z_1\leqslant z,\delta_1=1]|\\
&+|E\int_0^z C^{-3}(u)I[T_1\leqslant u\leqslant Z_1]\mathrm{d}W_1(u)|\\
&+|EC^{-2}(Z_1)I[Z_1\leqslant z]I[T_2\leqslant Z_1\leqslant Z_2]\delta_1|\\
&+|E\int_0^z C^{-3}(u)I[T_1\leqslant u\leqslant Z_1]I[T_2\leqslant u\leqslant Z_2]\mathrm{d}W_1(u)|\\
\leqslant &4\sigma_1^2
\end{aligned}$$

及

$$\begin{aligned}
&E|\psi((Z_i,\delta_i,T_i),(Z_i,T_i,\delta_i))\\
\leqslant &2(E\psi_1((Z_1,Z_1),(T_1,T_1))\delta_1)^2 + 2(E\psi_2((Z_1,Z_1),(T_1,T_1)))^2\\
\leqslant &4\beta_5(z), \quad (1.2.30)
\end{aligned}$$

由(1.2.30),得

$$\begin{aligned}
&P(\sqrt{n}\sigma^{-1}|R_{2n}| > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}})\\
\leqslant &P(\sqrt{n}\sigma_1^{-1}\frac{1}{n^2}|\sum_{i=1}^{n}(\psi((Z_i,T_i,\delta_i),(Z_i,T_i,\delta_i))\\
&-E\psi((Z_i,T_i,\delta_i),(Z_i,T_i,\delta_i))| > \alpha_z n^{-\frac{1}{2}})\\
\leqslant &\alpha_z n^{-1}E|\psi((Z_1,T_1,\delta_1),(Z_1,T_1,\delta_1))|^2
\end{aligned}$$

$$\leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}. \tag{1.2.31}$$

这证明了 $i=2$ 的情形.

类似地,我们能证 $i=6$ 的情形,即

$$P(\sqrt{n}\sigma^{-1}|R_{6n}| > \alpha\sigma_1^{-3}\beta_{14}(Z)n^{-\frac{1}{2}}) \leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}. \tag{1.2.32}$$

为证 $i=3$ 的情形,首先让我们作下面变换:

$$V_i = \begin{cases} W_1(Z_i), & 若\ \delta_i = 1, \\ 1, & 若\ \delta_i = 0. \end{cases} \tag{1.2.33}$$

易见 $V_1, V_2, \cdots, V_n \overset{i.i.d}{\sim} U[0,1]$,事实上

$$P(V_i \leqslant x) = P(W_1(Z_i) \leqslant x, \delta_i = 1) + P(1 \leqslant x, \delta_i = 0)$$

$$= \begin{cases} W_1(W_1^{-1}(x)), & 0 \leqslant x < 1, \\ P(\delta_i = 1) + P(\delta_i = 0), & x = 1, \end{cases}$$

其中 $W_1^{-1}(x) = \inf\{t: W_1(t) \geqslant x\}$. 既然 $W_1$ 连续,我们有

$$P(V_i \leqslant x) = x, \quad 0 \leqslant x \leqslant 1, \quad i = 1,2,\cdots,n.$$

设 $U_n(x) = n^{-1}\sum_{i=1}^{n} I[V_i \leqslant x]$ 是对应的经验分布函数,于是使用 Gijbels 与 Wang[37] 中引理 3 的证法,可得

$$R_{3n} = (1 - F(z)) \times \int_0^\infty\int_0^\infty\int_0^\infty k(u,v,w)$$

$$\mathrm{d}(U_n(u) - U(u))\mathrm{d}(C_n(v) - C(v))\mathrm{d}(C_n(w) - C(w)), \tag{1.2.34}$$

且 $0 < k(u,v,w) \leqslant C^{-3}(a_W)$. 应用 Gijbels 与 Wang[37] 中的引理 1,得

$$ER_{3n}^2(z) = O(n^{-3}). \tag{1.2.35}$$

因此,由 Tcbebychev 不等式与(1.2.35),我们有

$$P(\sqrt{n}\sigma^{-1}|R_{3n}(z)| > \alpha\sigma^{-3}\beta_{14}(z)n^{-\frac{1}{2}})$$

$$\leqslant \alpha_z n^{-1} \leqslant \alpha\sigma_1^{-3}\beta_{14}n^{-\frac{1}{2}}. \tag{1.2.36}$$

对 $i=4$,证明类似于 Gijbels 与 Wang[37] 中的引理 2 的证明.设

$$S_n(z)=\left\{\left|\frac{C_n(Z_i)-C(Z_i)}{C(Z_i)}\right|I[0<Z_i\leqslant z]<\frac{1}{2},i=1,2,\cdots,n\right\},$$

$S_n^c(z)$ 是 $S_n(z)$ 的补. 既然 $C(t)\geqslant C(a_W)$,于是由 Gijbels 与 Wang[37] 我们有

$$P(S_n^c(z))\leqslant 2n\mathrm{e}^{-\alpha n},\tag{1.2.37}$$

易见

$$\begin{aligned}&P(\sqrt{n}\sigma^{-1}|R_{4n}|>\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}})\\ \leqslant{}&P(\sqrt{n}\sigma^{-1}|R_{4n}|>\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}},S_n(z))+P(S_n^c(z)).\end{aligned}\tag{1.2.38}$$

注意到

$$I[0<Z_i\leqslant z]=I[a_W\leqslant Z_i\leqslant z],$$

因此,由不等式 $Q(x)\leqslant 2x^2,|x|<\frac{1}{2}$,容易看出在(1.2.38) 中第一项不超过

$$\begin{aligned}&P\left\{\frac{2\sqrt{n}}{n}\left|\sum_{i=1}^{n}\frac{(C_n(Z_i)-C(Z_i))^3}{C^4(Z_i)}I[a_W\leqslant Z_i\leqslant Z]\right|\right.\\ &\quad\left.>\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}\right\}\\ \leqslant{}&P\left\{\sup_x|C_n(x)-C(x)|>\sqrt[3]{\frac{\alpha\sigma_1^{-3}\beta_{14}(z)C^4(a_W)}{2n}}\right\}\\ \leqslant{}&\alpha\mathrm{e}^{-\alpha_z n^{1/3}},\end{aligned}\tag{1.2.39}$$

上面最后一个不等式由 Dvoretzky,Kiefer 与 Wolfowice 的定理得到.

易见,(1.2.37) - (1.2.39) 一起就证明了

$$P(\sqrt{n}\sigma^{-1}|R_{4n}|>\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}})\leqslant\alpha\sigma_1^{-3}\beta_{14}(z)n^{-1/2}.\tag{1.2.40}$$

由 Gijbels 与 Wang[37] 中的引理 4,我们有

$$P(\sqrt{n}\sigma^{-1}|R_{5n}| > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}) \leqslant \alpha e^{-\alpha n} < \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}. \tag{1.2.41}$$

由引理 1.2.1(i),得

$$\begin{aligned}
& P\left(\sqrt{n}\sigma^{-1}|R_{7n}| > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}\right) \\
\leqslant\ & P\left(\left|\frac{1}{n}\sum_{i=1}^{n}\zeta(Z_i,\delta_i,T_i)\right| > \sqrt{\alpha\sigma_1^{-2}\beta_{14}(z)}(\log n)^{-2}\right) \\
& + P\left(n|r_n| > \sqrt{\alpha\sigma_1^{-2}\beta_{14}(z)}(\log n)^2\right) \\
\leqslant\ & \alpha_z\frac{1}{n^2}n(\log n)^4 + 2ne^{\alpha n} \\
& + K\left[e^{-\alpha_z(\log n)^2} + \left(\frac{\alpha_z(\log n)^2}{5}\right)^{-2n} + e^{-\alpha_z(n(\log n)^2)^{1/2}} + e^{-\alpha_z n(\log n)^4}\right] \\
\leqslant\ & \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}},
\end{aligned} \tag{1.2.42}$$

且

$$\begin{aligned}
& P(\sqrt{n}\sigma^{-1}|R_{8n}| > \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}) \\
\leqslant\ & P(n|r_n| > \sqrt{2\alpha\sigma_1^{-2}\beta_{14}(z)}n^{\frac{1}{2}}) \\
\leqslant\ & \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}.
\end{aligned} \tag{1.2.43}$$

剩下还要证 $i=9$ 的情形. 设 $\widetilde{U}_n$ 定义(1.2.11)右边第一项,显然 $\widetilde{U}_n$ 是 $U$ 统计量. 根据引理 1.2.3 的证明,可证存在常数 $\alpha$ 使得

$$\sup_x|P(\sqrt{n}\sigma^{-1}\widetilde{U}_n \leqslant x) - \Phi(x)| \leqslant \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}. \tag{1.2.44}$$

因此,由(1.2.39),(1.2.29),(1.2.31),(1.2.36),(1.2.40)与引理 1.2.2,我们有

$$\begin{aligned}
& \sup_x|P(\sqrt{n}\sigma_1^{-1}(\Lambda_n(z) - \Lambda(z)) \leqslant x) - \Phi(x)| \\
\leqslant\ & \alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}.
\end{aligned} \tag{1.2.45}$$

既然

$$\begin{aligned}&-(\log(1-\hat{F}_n(z))+\Lambda(z))\\&=(\Lambda_n(z)-\Lambda(z))-(\log(1-\hat{F}_n(z))+\Lambda_n(z))\\&=\Lambda_n(z)-\Lambda(z)+\frac{R_{5n}}{1-F(z)},\end{aligned}\tag{1.2.46}$$

由(1.2.41),(1.2.45),(1.2.46)与引理1.2.2,我们得

$$\begin{aligned}&\sup_x|P(\sqrt{n}\sigma_1^{-1}(-\log(1-\hat{F}_n(z)-\Lambda(z))\leqslant x)-\Phi(x))\\&\leqslant\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}.\end{aligned}\tag{1.2.47}$$

因此

$$\begin{aligned}&P(\sqrt{n}\sigma^{-1}|R_{9n}|>\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}})\\&\leqslant P(\sqrt{n}\sigma_1^{-1}|\log(1-\hat{F}_n(z))+\Lambda(z)|^3>\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}},\\&\quad|\log(1-\hat{F}_n(z))+\Lambda(z)|<1)\\&\quad+P(|\log(1-\hat{F}_n(z))+\Lambda(z)|>1)\stackrel{\mathrm{d}}{=}P_1+P_2.\end{aligned}\tag{1.2.48}$$

由 Gijbels 与 Wang[37] 中的(3.14),知

$$P_2\leqslant\alpha\left[\mathrm{e}^{-\lambda n}+\left(\frac{n}{25}\right)^{-2n}\right],\tag{1.2.49}$$

其中 $\lambda$ 是某正常数.

又由(1.2.47)与不等式 $1-\Phi(x)\leqslant\frac{1}{\sqrt{2\pi}x}\mathrm{e}^{-\frac{x^2}{2}},x>0$ 可推得

$$\begin{aligned}P_1&\leqslant P(\sqrt{n}|\log(1-\hat{F}_n(z))+\Lambda(z)|>\alpha_z n^{\frac{1}{6}})\\&\leqslant|P(\sqrt{n}(-\log(1-\hat{F}_n(z))-\Lambda(z))>\alpha_z n^{\frac{1}{6}})-(1-\Phi(\alpha_z n^{\frac{1}{6}}))|\\&\quad+|P(\sqrt{n}(-\log(1-\hat{F}_n(z))-\Lambda(z))<-\alpha_z n^{\frac{1}{6}})-\Phi(-\alpha_z n^{\frac{1}{6}})|\\&\quad+2(1-\Phi(\alpha_z n^{\frac{1}{6}})\leqslant\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}},\end{aligned}\tag{1.2.50}$$

(1.2.48),(1.2.49)与(1.2.50)一起证明了

$$P(\sqrt{n}\sigma^{-1}|R_{9n}|>\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}})\leqslant\alpha\sigma_1^{-3}\beta_{14}(z)n^{-\frac{1}{2}}.$$

至此,引理1.2.4得证.

# §1.3 二元乘积限估计的 Berry-Essen 不等式

## 1.3.1 记号与主要结果

设$(X_i^0, Y_i^0), 1 \leqslant i \leqslant n$ 是 iid 非负的随机向量，具有共同的连续生存分布函数 $S(s,t) = P(X_1^0 > s, Y_1^0 > t)$. 设$(C_i, D_i), 1 \leqslant i \leqslant n$ 是对应于$(X_i^0, Y_i^0)$ 的删失时间向量且$(C_i, D_i), 1 \leqslant i \leqslant n$ iid 具有连续的生存分布函数 $G(s,t) = P(C_1 > s, D_1 > t)$. 设$(X_i^0, Y_i^0)$ 与$(C_i, D_i)$ 相互独立，$i = 1,2,\cdots,n$. 在二元随机删失模型中，我们观察到的数据为$(X_i, Y_i, \delta_i, \Delta_i), i = 1,2,\cdots,n$. 其中

$$X_i = \min(X_i^0, C_i), \qquad Y_i = \min(Y_i^0, D_i),$$

$$\delta_i = \begin{cases} 1, & 若\ X_i^0 \leqslant C_i, \\ 0, & 若\ X_i^0 > C_i, \end{cases} \qquad \Delta_i = \begin{cases} 1, & 若\ Y_i^0 \leqslant D_i, \\ 0, & 若\ Y_i > D_i. \end{cases}$$

使用上面的删失数据，Campbell 与 Földes[15] 定义了 $S(s,t)$ 的乘积限估计

$$\hat{S}_n(s,t) = \begin{cases} \prod_{i=1}^n \left[\frac{N_n(X_i,0)}{N_n(X_i,0)+1}\right]^{\alpha_i(s,0)} \prod_{j=1}^n \left[\frac{N_n(s,Y_j)}{N_n(s,Y_j)+1}\right]^{\beta_j(s,t)}, & 若\ N_n(s,t) > 0, \\ 0, & 其他, \end{cases}$$

其中

$$N_n(s,t) = \sum_{i=1}^n I[X_i > s, Y_i > t],$$

$$\alpha_i(s,t) = I[X_i \leqslant s, Y_i > t, \delta_i = 1], 1 \leqslant i \leqslant n,$$

$$\beta_j(s,t) = I[X_j > s, Y_j \leqslant t, \Delta_j = 1], 1 \leqslant j \leqslant n,$$

$I[\cdot]$ 是某事件的示性函数. 设

$$H(s,t) = P(X > s, Y > t),$$

$$H_{1x}(s,t) = P(X > s, Y > t, \delta = 1),$$

$$H(t|s) = P(Y > t|X > s),$$
$$H_{1y}(t|s) = P(Y > t, \Delta = 1|X > s),$$
$$\sigma_0^2 = -\int_0^s H^{-2}(u,0)\mathrm{d}H_{1x}(u,0) - \int_0^t H^{-2}(v|s)\mathrm{d}H_{1y}(v|s),$$

且约定 $\alpha$ 可表示任何所需的常数，$\alpha_0$ 可表示任何所需的绝对常数，在不同的地方它们均可表示不同值.

易见，若 $N_n(s,t)=0$，则 $-\log S_n(s,t)=\infty$. 注意 $P(N_n(s,t)=0)>0$，因此，以正概率 $-\log S_n(s,t)=\infty$. 为克服这一困难，我们将 $\Omega$ 分成两部分：

$$\Omega_0^{(n)} = \{\omega \in \Omega:\ N_n(s,t) > 0\},$$

且

$$\Omega_1^{(n)} = \Omega - \Omega_0^{(n)}.$$

既然

$$\begin{aligned} P(\Omega_1^{(n)}) &= P(N_n(s,t)=0) \\ &= \prod_{i=1}^n P(X_i \leqslant s \text{ 或 } Y_i \leqslant t) \\ &= (1-H(s,t))^n \leqslant \alpha n^{-k}, \end{aligned}$$

对任何使得 $H(s,t)>0$ 的 $s,t>0$ 与 $k>0$ 成立. 因此，我们可在 $\Omega_0^{(n)}$ 上讨论我们的问题.

设

$$A_1(x) = -H^{-1}(x,0)I[0 \leqslant x \leqslant s],$$
$$A_2(x) = -\int_0^{s\wedge x} H^{-2}(u,0)\mathrm{d}H_{1x}(u,0),$$
$$\tilde{g}(X_1,\delta_1) = A_1(X_1)\delta_1 + A_2(X_1),$$
$$B_1(x,y) = H^{-2}(x,0)I[0 \leqslant x \leqslant s, x < y],$$
$$B_2(x,y) = \int_0^{s\wedge x\wedge y} H^{-3}(u,0)\mathrm{d}H_{1x}(u,0),$$
$$\begin{aligned} \tilde{\psi}(X_1,\delta_1;X_2,\delta_2) &= B_1(X_1,X_2)\delta_1 + B_2(X_1,X_2) \\ &\quad - E[B_1(X_1,X_2)\delta_1 + B_2(X_1,X_2)\,|\,X_1,\delta_1], \end{aligned}$$

$$\tilde{h}(X_1,\delta_1;X_2,\delta_2) = \tilde{g}(X_1,\delta_1) + \tilde{g}(X_2,\delta_2) + \tilde{\psi}(X_1,\delta_1;X_2,\delta_2) + \tilde{\psi}(X_2,\delta_2;X_1,\delta_1),$$

$$H_{1y}(s,t) = P(X > s, Y > t, \Delta = 1),$$

$$H_{11}(s,t) = P(X > s, Y > t, \delta = 1, \Delta = 1),$$

$$m_s = \sum_{i=1}^{n} I[X_i > s],$$

$$H_n(t|s) = m_s^{-1}\sum_{i=1}^{n} I[X_i > s, Y_i > t],$$

$$H_{1yn}(t|s) = m_s^{-1}\sum_{i=1}^{n} I[X_i > s, Y_i > t, \Delta_i = 1],$$

$$A_{1s}(y) = -H^{-1}(y|s)I[0 \leqslant y \leqslant t],$$

$$A_{2s}(y) = -\int_0^{t\wedge y} H^{-2}(v|s)\mathrm{d}H_{1y}(v|s),$$

$$\tilde{g}_s(Y_1,\Delta_1) = A_{1s}(Y_1)\Delta_1 + A_{2s}(Y_2),$$

$$\widetilde{B}_{1s}(u,v) = H^{-2}(u|s)I[0 \leqslant u \leqslant t, u < v],$$

$$\widetilde{B}_{2s}(u,v) = \int_0^{t\wedge u\wedge v} H^{-3}(x|s)\mathrm{d}H_1(x|s),$$

$$\tilde{\psi}_s(Y_1,\Delta_1;Y_2,\Delta_2) = \widetilde{B}_{1s}(Y_1,Y_2)\Delta_1 + \widetilde{B}_{2s}(Y_1,Y_2) - E[\widetilde{B}_{1s}(Y_1,Y_2)\Delta_1 + \widetilde{B}_{2s}(Y_1,Y_2) \mid Y_1,\Delta_1],$$

$$\tilde{h}_s(Y_1,\Delta_1;Y_2,\Delta_2) = \tilde{g}_s(Y_1,\Delta_1) + \tilde{g}_s(Y_2,\Delta_2) + \tilde{\psi}_s(Y_1,\Delta_1;Y_2,\delta_2) + \tilde{\psi}_s(Y_2,\Delta_2;Y_1,\Delta_1),$$

$$h(X_i,\delta_i,Y_i,\Delta_i;X_j,\delta_j,Y_j,\Delta_j) = \tilde{h}(X_i,\delta_i;X_j,\delta_j) + \frac{\tilde{h}_s(Y_i,\Delta_i;Y_j,\Delta_j)I[X_i > s]I[X_j > s]}{H^2(s,0)}$$
$$- 2\left(\frac{\tilde{g}_s(Y_i,\Delta_i)I[X_i > s](I[X_j > s] - EI[X_j > s])}{H^2(s,0)} + \frac{\tilde{g}_s(Y_j,\Delta_j)I[X_j > s](I[X_i > s] - EI[X_i > s])}{H^2(s,0)}\right),$$

$$U_n = n^{-2}\sum_{i<j}h(X_i,\delta_i,Y_i,\Delta_i;X_j,\delta_j,Y_j,\Delta_j).$$

**定理 1.3.1** 对任何 $\omega \in \Omega_0^{(n)}$ 与任何使得 $H(s,t) > 0$ 的 $s,t > 0$,我们有

$$\log S_n(s,t) - \log S(s,t) = U_n + \frac{1}{2}n^{-1}\sigma_0^2 + R_n,$$

且

$$P(\sqrt{n}\sigma_0 | R_n | > \alpha(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}),$$

其中 $\sigma_0$ 如前所定义.

**定理 1.3.2** 对任何使得 $H(s,t) > 0$ 的 $s,t > 0$,存在绝对常数 $\alpha_0$ 使得

$$\sup_x |P(\sqrt{n}\sigma^{-1}(S_n(s,t) - S(s,t)) \leqslant x) - \Phi(x)| \leqslant \alpha_0 B n^{-\frac{1}{2}},$$

其中 $\Phi(x)$ 是标准正态分布函数,

$$B = H^{-3}(s,t)\left(\frac{1}{\sigma_0^3} + \frac{1}{\sigma_0^2} + \frac{1}{\sigma_0}\right),$$

$$\sigma^2 = S^2(s,t)\sigma_0^2.$$

### 1.3.2 定理的证明

设

$$Q(x) = \frac{1}{1+x} - (1-x), |x| < 1,$$

$$\tilde{\sigma}_1^2 = -\int_0^s H^{-2}(u,0)\mathrm{d}H_{1X}(u,0), \tilde{\sigma}_{1s}$$

$$= -\int_0^t H^{-2}(v|s)\mathrm{d}H_{1y}(v|s).$$

**引理 1.3.1** 对任何 $\omega \in \Omega_0^{(n)}$ 和使得 $H(s,t) > 0$ 的 $s,t > 0$,有

$$\log\hat{S}_n(s,t) - \log S(s,t)$$

$$= U_n + R_{1ns} + R_{2ns} + \sum_{i=1}^{5} R_{in}(t|s)$$

$$+\sum_{i=2}^{4}U_{n21}^{(i)}(t|s)+\sum_{i=2}^{4}U_{n2i}(t|s), \qquad (1.3.1)$$

且 $R_{1ns}$ 满足

$$P(\sqrt{n}\sigma_0^{-1}|\widetilde{R}_{1ns}|>\alpha(n\log n)^{-\frac{1}{2}})=o(n^{-\frac{1}{2}}),$$

而

$$R_{2ns}=-\frac{1}{2}n^{-1}\tilde{\sigma}_1^2,$$

$$R_{1n}(t|s)=\int_0^t H^{-3}(v|s)(H_n(v|s)$$
$$-H(v|s)^2\mathrm{d}(H_{1yn}(v|s)-H_{1y}(v|s)),$$

$$R_{2n}(t|s)=\sum_{k=3}^{\infty}(-1)^{k+1}\int_0^t H^{-(k+1)}(v|s)(H_n(v|s)$$
$$-H(v|s))^k\mathrm{d}H_{1yn}(v|s),$$

$$R_{3n}(t|s)=\sum_{k=3}^{\infty}\Big[(-1)^k k^{-1}\sum_{k=2}^{k}n^{-i+1}\binom{k}{i}$$
$$\times\int_0^t H^{-k}(v|s)(H_n(v|s)-H(v|s))^{k-i}\mathrm{d}H_{1yn}(v|s)\Big],$$

$$R_{4n}(t|s)=n^{-2}\sum_{i=1}^{n}\{2\tilde{g}_s(Y_i,\Delta_i)+\widetilde{B}_{2s}(Y_i,Y_i)$$
$$+\frac{1}{2}H^{-2}((Y_i|s)I[0\leqslant Y_i\leqslant t,\Delta_i=1]\}$$
$$-E\{2g(Z_i,\delta_i)+\widetilde{B}_{2s}(Y_i,Y_i)$$
$$+\frac{1}{2}H^{-2}(Y_i|s)I[0\leqslant Y_i\leqslant t,\Delta_i=1]\},$$

$$R_{5n}(t|s)=-\frac{1}{2}n^{-1}\tilde{\sigma}_{1s}^2,$$

$$U_{n21}^{(2)}(t|s)$$
$$=-\frac{1}{n^2}\sum_{i=1}^{n}\frac{\tilde{g}_s(Y_i,\Delta_i)I[X_i>s](I[X_i>s]-EI[X_i>s])}{H^2(s,0)},$$

$$U_{n21}^{(3)}(t|s) = \frac{1}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s]\,\frac{(H_n(s,0) - H(s,0))^2}{H^3(s,0)},$$

$$U_{n214}^{(4)}(t|s) = \frac{1}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s]$$

$$\cdot\frac{H(s,0) - H_n(s,0)}{H^2(s,0)}Q\left(\frac{H_n(s,0) - H(s,0)}{H(s,0)}\right),$$

$$U_{n22}(t|s) = -\frac{1}{n^2}\Big\{\sum_{i=1}^{n}(\tilde{g}_s(Y_i,\Delta_i)$$

$$+\tilde{\psi}(Y_i,\Delta_1;Y_i,\Delta_i)I[X_i > s]\Big\}\frac{H^2(s,0) - H_n^2(s,0)}{H_n^2(s,0)H^2(s,0)},$$

$$U_{n23}(t|s) = -\frac{H^2(s,0) - H_n^2(s,0)}{H^2(s,0)}$$

$$\cdot\int_0^t\frac{(H_n(v|s) - H(v|s))\mathrm{d}(H_{1yn}(v|s) - H_{1y}(v|s))}{H^2(v|s)},$$

$$U_{n24}(t|s) = \frac{H^2(s,0) - H_n^2(s,0)}{H^2(s,0)}$$

$$\cdot\int_0^t\frac{(H_n(v|s) - H(v|s))^2}{H^2(v|s)}\mathrm{d}H_{1y}(v|s).$$

证明：易见

$$\log\hat{S}_n(s,t) - \log S(s,t)$$
$$= (\log\hat{S}_n(s,0) - \log S(s,0)) + (\log\hat{S}_n(t|s) - \log S(t|s)). \tag{1.3.2}$$

由 Chang[17]，对 $\omega\in\Omega_0^{(n)}$，我们有

$$\log\hat{S}_n(s,0) - \log S(s,0)$$
$$= \frac{1}{n^2}\sum_{i<j}\tilde{h}(X_i,\delta_i;X_j,\delta_j) + \widetilde{R}_{1ns} + \widetilde{R}_{2ns}, \tag{1.3.3}$$

其中 $\widetilde{R}_{1ns}$ 满足

$$P(\sqrt{n}\tilde{\sigma}_1^{-1}|\widetilde{R}_{1ns}| > \alpha(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}), \tag{1.3.4}$$

且

$$\widetilde{R}_{2ns} = -\frac{1}{2}n^{-1}\tilde{\sigma}_1^2. \tag{1.3.5}$$

既然 $\sigma_0^2 = \tilde{\sigma}_1^2 + \tilde{\sigma}_{1s}^2$,用 $\sigma_0$ 取代(1.3.4)中 $\tilde{\sigma}_1$,(1.3.4)仍成立.

再仿效 Chang[17] 中定理 1 的证明,可得

$$\begin{aligned} &\log\hat{S}_n(t|s) - \log S(t|s) \\ &= \sum_{i<j}\frac{\tilde{h}_s(Y_i,\Delta_i;Y_j,\Delta_j)I[X_1 > s]I[X_j > s]}{m_s^2} \\ &\quad + \sum_{i=1}^{5}R_{in}(t|s), \end{aligned} \tag{1.3.6}$$

其中 $R_{in}(t|s)$ 如引理 1.3.1 所定义,$i = 1,2,\cdots,n$.

易见

$$\begin{aligned} &\sum_{i<j}\frac{\tilde{h}_s(Y_i,\Delta_i;Y_j,\Delta_j)I[X_i > s]I[X_j > s]}{m_s^2} \\ &= \frac{1}{n^2}\sum_{i<j}\frac{\tilde{h}_s(Y_i,\Delta_i;Y_j,\Delta_j)I[X_i > s]I[X_j > s]}{H^2(s,0)} \\ &\quad + \frac{1}{n^2}\sum_{i<j}\tilde{h}_s(Y_i,\Delta_i;Y_j,\Delta_j)I[X_i > s]I[X_j > s] \\ &\quad \cdot\frac{H^2(s,0) - H_n^2(s,0)}{H_n^2(s,0)H^2(s,0)} \\ &\stackrel{\mathrm{d}}{=} U_{n1}(t|s) + U_{n2}(t|s). \end{aligned} \tag{1.3.7}$$

注意到

$$\begin{aligned} &\frac{1}{n^2}\sum_{i<j}\tilde{h}_s(Y_i,\Delta_i;Y_j,\Delta_j)I[X_i > s]I[X_j > s] \\ &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s]I[X_j > s] \\ &\quad - \frac{1}{n^2}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s] \\ &\quad + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\psi}_s(Y_i,\Delta_i;Y_j,\Delta_j)I[X_i > s]I[X_j > s] \end{aligned}$$

$$-\frac{1}{n^2}\sum_{i=1}^{n}\tilde{\psi}_s(Y_i,\Delta_i;Y_i,\Delta_i)I[X_i>s], \qquad (1.3.8)$$

且

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\psi}_s(Y_i,\Delta_i;Y_j,\Delta_j)I[X_i>s]I[X_j>s]$$

$$=-H_n^2(s,0)\int_0^t\frac{H_n(v|s)-H(v|s))\mathrm{d}(H_{1yn}(v|s)-H_{1y}(v|s))}{H^2(v|s)}$$

$$+H_n^2(s,0)\int_0^t\frac{(H_n(v|s)-H(v|s))^2\mathrm{d}H_{1y}(v|s)}{H^2(v|s)}. \qquad (1.3.9)$$

由(1.3.8)与(1.3.9),可得

$$\widetilde{U}_{n2}(t|s)=\frac{1}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i>s]\frac{H^2(s,0)-H_n^2(s,0)}{H_n(s,0)H^2(s,0)}$$

$$-\frac{1}{n^2}\Big\{\sum_{i=1}^{n}(\tilde{g}_s(Y_i,\Delta_i)+\tilde{\psi}(Y_i,\Delta_i;\ Y_i,\Delta_i)I[X_i>s]\Big\}$$

$$\times\frac{H^2(s,0)-H_n^2(s,0)}{H_n^2(s,0)H^2(s,0)}-\frac{H^2(s,0)-H_n^2(s,0)}{H^2(s,0)}$$

$$\int_0^t\frac{(H_n(v|s)-H(v|s))\mathrm{d}(H_{1yn}(v|s)-H_{1y}(v|s))}{H^2(v|s)}$$

$$+\frac{H^2(s,0)-H_n^2(s,0)}{H^2(s,0)}\int_0^t\frac{(H_n(v|s)-H(v|s))^2}{H^2(v|s)}\mathrm{d}H_{1y}(v|s)$$

$$\stackrel{\mathrm{d}}{=}U_{n21}(t|s)+U_{n22}(t|s)+U_{n23}(t|s)+U_{n24}(t|s). \qquad (1.3.10)$$

设 $Q(x)$ 如前所定义,我们有

$$U_{n21}(t|s)=\frac{1}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i>s]\frac{H(s,0)-H_n(s,0)}{H^2(s,0)}$$

$$\times\left[1-\frac{H_n(s,0)-H(s,0)}{H(s,0)}+Q\left(\frac{H_n(s,0)-H(s,0)}{H(s,0)}\right)\right]$$

$$+\frac{1}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i>s]\frac{(H(s,0)-H_n(s,0))}{H^2(s,0)}$$

$$
\begin{aligned}
&= -\frac{2}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s]\frac{(H_n(s,0)-H(s,0))}{H^2(s,0)} \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s]\frac{(H_n(s,0)-H(s,0))^2}{H^3(s,0)} \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s]\frac{(H(s,0)-H_n(s,0)}{H^2(s,0)} \\
&\quad \cdot Q\left(\frac{H_n(s,0)-H(s,0)}{H(s,0)}\right) \\
&= -\frac{2}{n^2}\sum_{i<j}\left[\frac{\tilde{g}_s(Y_i,\Delta_i)I[X_i > s](I[X_j > s]-EI[X_j > s])}{H^2(s,0)}\right. \\
&\quad \left. + \frac{\tilde{g}_s(Y_j,\Delta_j)I[X_j > s](I[X_i > s]-EI[X_i > s])}{H^2(s,0)}\right] \\
&\quad - \frac{2}{n^2}\sum_{i=1}^{n}\frac{\tilde{g}_s(Y_i,\Delta_i)I[X_i > s](I[X_i > s]-EI[X_i > s])}{H^2(s,0)} \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s]\frac{(H_n(s,0)-H(s,0))^2}{H^3(s,0)} \\
&\quad + \sum_{i=1}^{n}\tilde{g}_s(Y_i,\Delta_i)I[X_i > s]\frac{(H(s,0)-H_n(s,0))}{H^2(s,0)} \\
&\quad \cdot Q\left(\frac{H_n(s,0)-H(s,0)}{H(s,0)}\right) \\
&\overset{\mathrm{d}}{=} U_{n21}^{(1)}(t\mid s) + U_{n21}^{(2)} + U_{n21}^{(3)}(t\mid s) + U_{n21}^{(4)}(t\mid s). \qquad (1.3.11)
\end{aligned}
$$

综合(1.3.2) - (1.3.7),(1.3.10)与(1.3.11),引理1.3.1得证.

**引理1.3.2** 对任何 $\alpha > 0$,有

$$
\begin{aligned}
&P(\sqrt{n}\sigma_0^{-1}\mid R_{in}(t\mid s)\mid > \alpha(n\log n)^{-\frac{1}{2}}) \\
&= o(n^{\frac{1}{2}}), i = 1,2,\cdots,5. \\
&P(\sqrt{n}\sigma_0^{-1}\mid U_{n21}^{(i)}(t\mid s)\mid > \alpha(n\log n)^{-\frac{1}{2}}) \\
&= o(n^{-\frac{1}{2}}), i = 2,3,4.
\end{aligned}
$$

$$P(\sqrt{n}\sigma_0^{-1}|U_{n2i}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}})$$
$$= o(n^{-\frac{1}{2}}), i = 2,3,4.$$

我们将在 1.3.3 中证明该引理.

**引理 1.3.4** 对任何使得 $H(s,t) > 0$ 的 $s,t > 0$,存在绝对常数 $\alpha_0$ 使得

$$\sup_x |P(\sqrt{n}\sigma_0^{-1}U_n \leqslant x) - \Phi(x)| \leqslant \frac{\alpha_0}{H^3(s,t)}\left(\frac{1}{\sigma_0^3} + \frac{1}{\sigma_0^2}\right).$$

证明:分部积分且注意到$\frac{H(s,t)}{H(s,0)} = H(t|s)$ 这一事实,可得

$$E\frac{\tilde{g}_s(Y_1,\Delta_1)I[X_1 > s]}{H(s,0)}$$
$$= -\int_0^t H^{-1}(v|s)H^{-1}(s,0)\mathrm{d}_v H_{1y}(s,v)$$
$$- \frac{1}{H(s,0)}\int_0^t\left[\int_0^y H^{-2}(v|s)\mathrm{d}H_{1y}(v|s)\right]\mathrm{d}_y H(s,y)$$
$$- \frac{1}{H(s,0)}\int_t^\infty\int_0^t H^{-2}(v|s)\mathrm{d}H_{1y}(v|s)\mathrm{d}_y H(s,y)$$
$$= -\int_0^t H^{-1}(v|s)\mathrm{d}H_{1y}(v|s)$$
$$- \left(\int_0^t H^{-2}(v|s)\mathrm{d}H_{1y}(v|s)\right)\frac{H(s,t)}{H(s,0)}$$
$$+ \int_0^t H^{-2}(v|s)\,\frac{H(s,v)}{H(s,0)}\mathrm{d}H_{1y}(v|s)$$
$$+ \left(\int_0^t H^{-2}(v|s)\mathrm{d}H_{1y}(v|s)\right)\frac{H(s,t)}{H(s,0)} = 0, \qquad (1.3.12)$$

其中 $\mathrm{d}_v H_{1y}(s,v)$ 与 $\mathrm{d}_y H(s,y)$ 分别定义对固定的 $s$ 关于 $v$ 与 $y$ 进行 Lebesque-Stieljes 积分. 下面类似的记号均按此理解.

由(1.3.12),得

$$E\frac{\tilde{h}_s(Y_1,\Delta_1;Y_2,\Delta_2)I[X_1 > s]I[X_2 > s]}{H^2(s,0)} = 0. \qquad (1.3.13)$$

Chang[17] 证明了

$$E\tilde{h}(X_1,\delta_1;X_2,\delta_2)=0. \tag{1.3.14}$$

因此,由(1.3.13),(1.3.14)我们有

$$Eh(X_1,\delta_1,Y_1,\delta_1;X_2,\delta_2,Y_2,\Delta_2)=0, \tag{1.3.15}$$

其中 $h$ 如前所定义. 易见 $U_n$ 是具有对称核$h$,且 $Eh=0$ 的二阶 $U$ 统计量.

注意到

$$E[\tilde{h}(X_1,\delta_1;X_2,\delta_2)\mid X_1,\delta_1]=\tilde{g}(X_1,\delta_1),$$

$$E\left[\frac{\tilde{h}_s(Y_1,\Delta_1;Y_2,\Delta_2)I[X_1>s]I[X_2>s]}{H^2(s,0)}\mid X_1,Y_1,\delta_1\right]$$

$$=\frac{\tilde{g}_s(Y_1,\Delta_1)I[X_1>s]}{H(s,0)},$$

$$E[\tilde{g}_s(Y_1,\Delta_1)I[X_1>0](I[X_2>s]-EI[X_2>s])]=0,$$

因此,我们有

$$\begin{aligned}&g(X_1,\delta_1;Y_1,\Delta_1)\\&\stackrel{d}{=}E[h(X_1,\delta_1,Y_1,\Delta_1;X_2,\delta_2,Y_2,\Delta_2)\mid X_1,\delta_1,Y_1,\Delta_1]\\&=\tilde{g}(X_1,\delta_1)+\frac{\tilde{g}_s(Y_1,\Delta_1)I[X_1>s]}{H(s,0)}.\end{aligned} \tag{1.3.16}$$

设

$$\begin{aligned}&\psi(X_1,\delta_1,Y_1,\Delta_1;X_2,\delta_2,Y_2,\Delta_2)\\&=\tilde{\psi}(X_1,\delta_1,X_2,\delta_2)+\tilde{\psi}(X_2,\delta_2,X_1,\delta_1)\\&\quad+\frac{\tilde{\psi}_s(Y_1,\Delta_1;Y_2,\Delta_2)I[X_1>s]I[X_2>s]}{H^2(s,0)}\\&\quad\cdot\frac{\tilde{\psi}_s(Y_2,\Delta_2;Y_1,\Delta_1)I[X_2>s]I[X_1>s]}{H^2(s,0)}\\&\quad-2\left(\frac{\tilde{g}_s(Y_1,\Delta_1)I[X_1>s](I[X_2>s]-EI[X_2>s])}{H^2(s,0)}\right.\end{aligned}$$

$$+\frac{\tilde{g}_s(Y_2,\Delta_2)I[X_2>s](I[X_1>s]-EI[X_1>s])}{H^2(s,0)}\Bigg).$$

(1.3.17)

因而 $U_n$ 能表示成

$$U_n=\frac{1}{n^2}\Big[(n-1)\sum_{i=1}^{n}g(X_i,\delta_i;Y_i,\Delta_i)$$

$$+\sum_{i<j}\psi(X_i,\delta_i,Y_i,\Delta_i;X_j,\delta_j,Y_j,\Delta_j)\Big].\quad(1.3.18)$$

由 Chang 与 Rao[16]

$$E\tilde{g}^2(X_1,\delta_1)=-\int_0^s H^{-2}(u,0)\mathrm{d}H_{1x}(u,0)=\tilde{\sigma}_1^2,$$

(1.3.19)

且直接计算可得

$$E\frac{\tilde{g}_s^2(Y_1,\Delta_1)I[X_1>s]}{H^2(s,0)}=-\int_0^t H^{-2}(v\,|\,s)\mathrm{d}H_{1y}(v\,|\,s)=\tilde{\sigma}_{1s}^2,$$

(1.3.20)

于是由(1.3.19),(1.3.20),(1.3.12)并注意应用下面事实:

$$\tilde{g}(X_1,\delta_1)I[X_1>s]=-\int_0^s H^{-2}(u,0)\mathrm{d}H_{1x}(u,0),$$

(1.3.21)

即可得

$$Eg^2(X_1,\delta_1,Y_1,\Delta_1)$$

$$=E\tilde{g}^2(X_1,\delta_1)+2E\frac{\tilde{g}(X_1,\delta_1)\tilde{g}_s(Y_1,\Delta_1)I[X_1>s]}{H(s,0)}$$

$$+\frac{E\tilde{g}_s^2(Y_1,\Delta_1)I[X_1>s]}{H^2(s,0)}=\sigma_1^2+\sigma_{1s}^2=\sigma_0^2.\quad(1.3.22)$$

又仿 Chang 与 Rao[16] 中(13)的证明,可得

$$E\,|\tilde{g}(X_1,\delta_1)|^3$$

$$\leqslant 7\int_0^s H^{-3}(u,0)\mathrm{d}H_{10}(u,0)+\frac{3}{2}\left(\int_0^t H^{-2}(u,0)\mathrm{d}H_{10}(u,0)\right)^2$$

$$\leqslant \alpha_0 H^{-3}(s,0). \tag{1.3.23}$$

类似地有

$$E\left|\frac{\tilde{g}_S(Y_1,\Delta_1)I[X_1>s]}{H(s,0)}\right|^3 \leqslant \alpha_0 H^{-3}(t|s). \tag{1.3.24}$$

由(1.3.24)及Hölder不等式得

$$E\left|\frac{\tilde{g}_S(Y_1,\Delta_1)I[X_1>s]}{H(s,0)}\right| \leqslant \alpha_0\ H^{-1}(t|s). \tag{1.3.25}$$

于是综合(1.3.23)－(1.3.25),及(1.3.21)与(1.3.20),即得

$$\begin{aligned}
&E|g(X_1,\delta_1,Y_1,\Delta_1)|^3\\
&= E|\tilde{g}(X_1,\delta_1)|^3\\
&\quad + 3E\left(\tilde{g}^2(X_1,\delta_1)I[X_1>s]\left|\frac{\tilde{g}_s(Y_1,\Delta_1)I[X_1>s]}{H(s,0)}\right|\right)\\
&\quad + 3E\left(|\tilde{g}(X_1,\delta_1)I[X_1>s]|\frac{\tilde{g}_s^2(Y_1,\Delta_1)I[X_1>s]}{H(s,0)}\right)\\
&\quad + E\left|\frac{\tilde{g}_s(Y_1,\Delta_1)I[X_1>s]}{H(s,0)}\right|^3\\
&\leqslant \alpha_0 H^{-3}(s,t).
\end{aligned} \tag{1.3.26}$$

由Chang与Rao[16],有

$$\begin{aligned}
&E\tilde{h}^2(X_1,\delta_1;\ X_2,\delta_2)\\
&\leqslant -\ \alpha_0\left(\int_0^s H^{-2}(u,0)\mathrm{d}H_{10}(u,0)+\int_0^s H^{-3}(u,0)\mathrm{d}H_{10}(u,0)\right)\\
&\leqslant \alpha_0\int_0^s H^{-3}(u,0)\mathrm{d}H_{10}(u,0),
\end{aligned} \tag{1.3.27}$$

类似计算可得

$$\begin{aligned}
&E\left[\frac{\tilde{h}_S(Y_1,\Delta_1;\ Y_2,\Delta_2)I[X>s]I[X_2>s]}{H^2(s,0)}\right]^2\\
&= E\left\{\frac{I[X_1>s]I[X_2>s]}{H^2(s,0)}E\left[\left(\frac{\tilde{h}_S(Y_1,\Delta_1;Y_2,\Delta_2)}{H(s,0)}\right)^2\middle|\ X_1,\delta_1;\ x_2,\delta_2\right]\right\}\\
&\leqslant -\ \alpha_0\int_0^t H^{-3}(v|s)\mathrm{d}H_{01}(v|s).
\end{aligned} \tag{1.3.28}$$

由$(X_1, Y_1, \delta_1)$与$X_2$的独立性及(1.3.20),得

$$E\left[\frac{\tilde{g}_s(Y_1,\Delta_1)I[X_1>0](I[X_2>s]-EI[X_2>s])}{H^2(s,0)}\right]^2$$
$$\leqslant H^{-3}(s,t).$$
(1.3.29)

于是由(1.3.27)-(1.3.29)及$h$的定义知

$$Eh^2(X_1,\delta_1,Y_1,\Delta_1;\ X_2,\delta_2,Y_2,\Delta_2)\leqslant \alpha_0 H^{-2}(s,t).$$
(1.3.30)

而由(1.3.22),(1.3.26)及(1.3.30),即得

$$\frac{E|g(X_1,\delta_1;\ Y_1,\Delta_2)|^3}{(Eg^2(X_1,\delta_1;Y_1,\Delta_1))^{\frac{3}{2}}}+\frac{Eh^2(X_1,\delta_1,Y_1,\Delta_1;\ X_2,\delta_2,Y_2,\Delta_2)}{Eg^2(X_1,\delta_1;\ Y_1,\Delta_1)}$$
$$\leqslant \frac{\alpha_0}{H^3(s,t)}\left(\frac{1}{\sigma_0^3}+\frac{1}{\sigma_0^2}\right). \tag{1.3.31}$$

记$\widetilde{B}=H^{-3}(\sigma_0^{-3}+\sigma_0^{-2})$,则存在绝对常数$\alpha_0$使得

$$\sup_x\left|P\left(\frac{n}{n-1}\sqrt{n}\sigma_0^{-1}U_n\leqslant x\right)-\Phi(x)\right|\leqslant \alpha_0\widetilde{B}n^{-\frac{1}{2}}.$$
(1.3.32)

由于

$$\text{Var } U_n=\frac{(n-1)(n-2)}{n^3}\sigma_0^2$$
$$+\frac{n-1}{2n^3}Eh^2(X_1,\delta_1,Y_1,\Delta_1;X_2,\delta_2,Y-2\Delta_2),$$

因而$n$充分大时

$$P\left(\left|\frac{\sqrt{n}}{n-1}\sigma_0^{-1}U_n\right|>\widetilde{B}n^{-\frac{1}{2}}\right)\leqslant \alpha_0\widetilde{B}n^{-\frac{1}{2}}. \tag{1.3.33}$$

于是由(1.3.31)-(1.3.33)及引理1.2.2知,存在绝对常数$\alpha_0$使得

$$\sup_x|P(\sqrt{n}\sigma^{-1}U_n\leqslant x)-\Phi(x)|\leqslant \alpha_0\widetilde{B}n^{-\frac{1}{2}}. \tag{1.3.34}$$

从而引理1.3.4得证.

**引理 1.3.5** 对任何使得 $H(s,t)>0$ 的 $s,t>0$,存在绝对常数 $\alpha_0$ 使得

$$\sup_x |P(\sqrt{n}\sigma_0^{-1}(\log\hat{S}_n(s,t)-\log S(s,t))\leqslant x)-\Phi(x)|$$
$$\leqslant \frac{\alpha_0}{H^3(s,t)}\left(\frac{1}{\sigma_0^3}+\frac{1}{\sigma_0^2}+\frac{1}{\sigma_0}\right). \tag{1.3.35}$$

证明:注意到 $\sigma_0\leqslant H^{-3}(s,t)\sigma_0^{-1}$,因而我们有

$$\sup_x |P(\sqrt{n}\sigma_0^{-1}(U_n+\frac{1}{2}n^{-1}\sigma_0^2)\leqslant x)-\Phi(x)|$$
$$\leqslant \sup_x |P(\sqrt{n}\sigma_0^{-1}U_n\leqslant x)-\Phi(x)|$$
$$+\sup_x\left|\Phi\left(x-\frac{n^{-\frac{1}{2}}\sigma_0}{2}\right)-\Phi(x)\right|$$
$$\leqslant \alpha_0\widetilde{B}n^{-\frac{1}{2}}+\frac{\sigma_0}{2\sqrt{2\pi}}n^{-\frac{1}{2}}$$
$$\leqslant \frac{\alpha_0}{H^3(s,t)}\left(\frac{1}{\sigma_0^3}+\frac{1}{\sigma_0^2}+\frac{1}{\sigma_0}\right). \tag{1.3.36}$$

应用定理 1.3.1 并再次应用引理 1.3.3,由(1.3.36) 即知引理 1.3.5 成立.

定理 1.3.2 的证明:仿 Chang[17] 中定理 3 的证明可证定理 2. 这里我们给出证定理 1.3.2 的梗概.

注意到

$$|\hat{S}_n(s,t)-S(s,t)|<1,$$

我们有

$$\sup_x |P(\sqrt{n}\sigma^{-1}(S_n(s,t)-S(s,t)\leqslant x)-\Phi(x)|$$
$$\leqslant \sup_{x<\sqrt{n}\sigma^{-1}}\left|P\left[\log\hat{S}_n(s,t)\leqslant\log\left(S(s,t)+\frac{\sigma x}{\sqrt{n}}\right)\right]-\Phi(x)\right|$$
$$+2(1-\Phi(\sigma_0^{-1}\sqrt{n}))$$
$$=\sup_{x<\sqrt{n}\sigma^{-1}}|P[\sqrt{n}\sigma_0^{-1}(\log\hat{S}_n(s,t)-\log S(s,t))$$

$$\leqslant \sqrt{n}\sigma_0^{-1}\log\left(1+\frac{\sigma_0 x}{\sqrt{n}}\right)\Big] - \Phi(x)| + \frac{2\sigma}{\sqrt{2\pi}}n^{-\frac{1}{2}}$$

$$\leqslant \alpha_0 Bn^{-\frac{1}{2}} + \sup_x|\Phi(\sqrt{n}\sigma_0^{-1}\log\left(1+\frac{\sigma_0 x}{\sqrt{n}}\right) - \Phi(x)| + \frac{2\sigma}{\sqrt{2\pi}}n^{-\frac{1}{2}}$$

$$\leqslant \alpha_0 Bn^{-\frac{1}{2}} + \sigma_0 n^{-\frac{1}{2}} + \frac{2\sigma}{\sqrt{2\pi}}n^{-\frac{1}{2}}$$

$$\leqslant \alpha_0 Bn^{-\frac{1}{2}}.$$

至此，定理 1.3.2 得证.

### 1.3.3 引理 1.3.2 的证明

根据 $H_n(v|s)$，$H_{1yn}(v|s)$，$H(v|s)$ 以及 $H_{1y}$ 的定义，我们有

$$R_{1n}(t|s) = \int_0^t H^{-3}(v|s)\left[\frac{H_n^2(s,v)(H_n(s,0)-H(s,0))^2}{H_n^2(s,0)H^2(s,0)}\right.$$

$$+\frac{2H_n(s,v)(H(s,0)-H_n(s,0))(H(s,v)-H_n(s,v))}{H_n(s,0)H^2(s,0)}$$

$$\left.+\frac{(H_n(s,v)-H(s,v))^2}{H^2(s,0)}\right]$$

$$\times \mathrm{d}_v\left(\frac{H_{01n}(s,v)(H(s,0)-H_n(s,0))}{H_n(s,0)H(s,0)}+\frac{H_{01n}(s,v)-H_{01}(s,v)}{H(s,0)}\right)$$

$$= -\frac{(H_n(s,0)-H(s,0))^3}{H_n^3(s,0)H^3(s,0)}\int_0^t H^{-3}(v|s)H_n^2(s,v)\mathrm{d}_v H_{01n}(s,v)$$

$$+\frac{2(H_n(s,0)-H(s,0))^2}{H_n^2(s,0)H^3(s,0)}$$

$$\cdot\int_0^t H^{-3}(v|s)H_n(s,v)(H_{01n}(s,v)-H_{01}(s,v))\mathrm{d}_v H_{01n}(s,v)$$

$$+\frac{H(s,0)-H_n(s,0)}{H_n(s,0)H^3(s,0)}$$

$$\cdot\int_0^t H^{-3}(v|s)(H_n(s,v)-H(s,v))^2\mathrm{d}_v H_{01n}(s,v)$$

$$
\begin{aligned}
&+\frac{(H_n(s,0)-H(s,0))^2}{H_n^2(s,0)H^3(s,0)} \\
&\cdot\int_0^t H^{-3}(v|s)H_n^2(s,v)\mathrm{d}_v(H_{01n}(s,v)-H_{01}(s,v)) \\
&+\frac{2(H(s,0)-H_n(s,0))}{H_n(s,0)H^3(s,0)} \\
&\times\int_0^t H^{-3}(v|s)H_n(s,v)(H_n(s,v)-H(s,v)) \\
&\cdot\mathrm{d}_v(H_{01n}(s,v)-H_{01}(s,v)) \\
&+\frac{1}{H^3(s,0)}\int_0^t H^{-3}(v|s)(H_n(s,v)-H(s,v))^2 \\
&\cdot\mathrm{d}_v(H_{01n}(s,v)-H_{01}(s,v)) \\
\stackrel{\mathrm{d}}{=}&\sum_{i=1}^{6}R_{1in}(t|s). \qquad (1.3.37)
\end{aligned}
$$

易见

$$
|R_{11n}(t|s)|\leqslant\frac{(H_n(s,0)-H(s,0))^3}{H^3(s,t)}. \qquad (1.3.38)
$$

由(1.3.37)与Dvoretzky, Kiefer与Wolfowitz定理(下面简称DKW定理),得

$$
\begin{aligned}
&P(\sqrt{n}\sigma_0^{-1}R_{11n}(t|s)|>(n\log n)^{-\frac{1}{2}}) \\
\leqslant&P(|H_n(s,0)-H(s,0)|>n^{-\frac{1}{3}}(\log n)^{-\frac{1}{6}}\sigma_0^{\frac{1}{3}}H(s,t)) \\
\leqslant&\mathrm{e}^{-2(n/\log n)^{\frac{1}{3}}\sigma^{\frac{1}{3}}H(s,t)}=o(n^{-\frac{1}{2}}). \qquad (1.3.39)
\end{aligned}
$$

既然

$$
\begin{aligned}
&|R_{12n}(t|s)| \\
\leqslant&\frac{2(H_n(s,0)-H(s,0))^2}{H^3(s,0)}\sup_{0\leqslant v\leqslant t}|H_{1yn}(s,v)-H_{1y}(s,v)|, \\
&R_{13n}(t|s)
\end{aligned}
$$

$$\leqslant \frac{|H_n(s,0) - H(s,0)|}{H^3(s,0)} \sup_{0 \leqslant v \leqslant t} |H_n(s,0) - H(s,0)|^2,$$

$$R_{14n}(t|s)$$

$$\leqslant \frac{(H_n(s,0) - H(s,0))^2}{H^3(s,t)} \sup_{0 \leqslant v \leqslant t} |H_{1yn}(s,v) - H_{1y}(s,v)|,$$

由 DKW 定理,类似(1.3.39) 的证明,我们有

$$P(\sqrt{n}\sigma_0^{-1}|R_{1in}(t|s)| > (n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}),$$
$$i = 2,3,4. \tag{1.3.40}$$

再由 DKW 定理,Tchebyshev 不等式及 Chang 与 Rao[16] 附录中引理,我们得

$$\begin{aligned}
&P(\sqrt{n}\sigma_0^{-1}|R_{15n}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) \\
&\leqslant 2P(|H_n(s,0) - H(s,0)| > \alpha H^4(s,t) n^{-\frac{1}{3}}, H_n(s,t) \\
&\quad \geqslant \frac{1}{2}H(s,t)) + 2P(H_n(s,t) < \frac{1}{2}H(s,0)) \\
&\quad + P\Big(\Big|\int_0^t (H_n(s,v) - H(s,v)) \\
&\quad \cdot \mathrm{d}_v(H_{1yn}(s,v) - H_{1y}(s,v)\Big| > n^{-\frac{2}{3}}(\log n)^{-\frac{1}{2}}\Big) \\
&\quad + P\Big(\Big|\int_0^t (H_n(s,v) - H(s,v))^2 \\
&\quad \cdot \mathrm{d}_v(H_{1yn}(s,v) - H_{1y}(s,v)\Big| > n^{-\frac{2}{3}}(\log n)^{-\frac{1}{2}}\Big) \\
&= 2\mathrm{e}^{-2\alpha H^8(s,t) n^{\frac{1}{3}}} + 2\mathrm{e}^{-\frac{H^2(s,t)}{2}n} + \alpha n^{-\frac{2}{3}}(\log n) + \alpha n^{-\frac{5}{3}}(\log n) \\
&= o(n^{-\frac{1}{2}}).
\end{aligned} \tag{1.3.41}$$

类似地,我们有

$$P(\sqrt{n}\sigma_0^{-1}|R_{16n}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}). \tag{1.3.42}$$

综合(1.3.37),(1.3.39) - (1.3.42),得

$$P(\sqrt{n}\sigma_0^{-1}|R_{1n}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}). \tag{1.3.43}$$

再参考 Chang[17] 中引理 4 在 $i = 3$ 时的证明,可证

$$\begin{aligned}&P(\sqrt{n}\sigma_0^{-1}|R_{2n}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}})\\ \leqslant\ &P(2\sup_{0\leqslant v\leqslant t}|H_n(v|s) - H(v|s)|^3H^{-4}(v|s) > \alpha n^{-1}(\log n)^{-\frac{1}{2}})\\ &+ P(\sup_{0\leqslant v\leqslant t}|H_n(v|s) - H(v|s)| > \frac{1}{2}H(v|s)).\end{aligned} \tag{1.3.44}$$

既然

$$\begin{aligned}&\sup_{0\leqslant v\leqslant t}|H_n(v|s) - H(v|s)|\\ \leqslant\ &\sup_{0\leqslant v\leqslant t}H^{-1}(s,0)(|H_n(s,0) - H(s,0)| + |H_n(s,v) - H(s,v)|),\end{aligned} \tag{1.3.45}$$

由(1.3.44),(1.3.45) 与 DKW 定理,有

$$P(\sqrt{n}\sigma_0^{-1}|R_{2n}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}). \tag{1.3.46}$$

类似地,参考 Chang[17] 中引理 4 在 $i = 3$ 时的证明,我们有

$$\begin{aligned}&P(\sqrt{n}\sigma_0^{-1}|R_{3n}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}})\\ \leqslant\ &P([\max(\sup_{0\leqslant v\leqslant t}|H_n(v|s) - H(v|s)|, n^{-\frac{1}{2}})]^3 2^3 H^{-3}(v|s)\\ &\sum_{k=0}^{\infty}2^k3^{-k} > \alpha n^{-1}(\log n)^{-\frac{1}{2}})\\ &+ P(\sup_{0\leqslant v\leqslant t}|H_n(v|s) - H(v|s)| > \frac{1}{3}H(v|s)),\end{aligned} \tag{1.3.47}$$

因此,由(1.3.45),(1.3.47) 与 DKW 定理,得

$$P(\sqrt{n}\sigma_0^{-1}|R_{3n}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}). \tag{1.3.48}$$

使用类似于 Chang[17] 中引理在 $i = 1$ 时的证法,可证

$$P(\sqrt{n}\sigma_0^{-1}|R_{4n}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}). \tag{1.3.49}$$

由 Tchebyshev 不等式

$$\begin{aligned} &P(\sqrt{n}\sigma_0^{-1}|U_{n21}^{(2)}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) \\ &\leqslant \alpha n^{-1}\log nE\frac{\tilde{g}_s^2(Y_i,\Delta_i)I[X_i>s](I[X_i>s]-EI[X_i>s])^2}{H^4(s,0)} \\ &= o(n^{-\frac{1}{2}}), \end{aligned} \tag{1.3.50}$$

且

$$\begin{aligned} &P(\sqrt{n}\sigma_0^{-1}|U_{n21}^{(3)}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) \\ &\leqslant P\left(\frac{1}{n}\left|\sum_{i=1}^n\frac{\tilde{g}_s(Y_i,\Delta_i)I[X_i>s]}{H(s,0)}\right| > \alpha n^{-\frac{1}{6}}\right) \\ &\quad + P\left(|H_n(s,0)-H(s,0)| > H(s,0)n^{-\frac{5}{12}}(\log n)^{-\frac{1}{4}}\right) \\ &= o(n^{-\frac{1}{2}}). \end{aligned} \tag{1.3.51}$$

由不等式 $Q(x)\leqslant 2x^2, |x|<\frac{1}{2}$ 并使用(1.3.51)同样的证法,可证

$$\begin{aligned} &P(\sqrt{n}\sigma_0^{-1}|U_{n21}^{(4)}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}, \\ &\quad |H_n(s,0)-H(s,0)| < \frac{1}{2}H(s,0)) \\ &\quad + P(|H_n(s,0)-H(s,0)| \geqslant \frac{1}{2}H(s,0)) \\ &\leqslant P\left(\sqrt{n}\sigma_0^{-1}\frac{1}{2}\left|\sum_{i=1}^n\tilde{g}_s(Y_i,\Delta_i)I[X_i>s]\right.\right. \\ &\quad \left.\frac{|H_n(s,0)-H(s,0)|^3}{H^4(s,0)} > \alpha(n\log n)^{-\frac{1}{2}}\right) \\ &\quad + o(n^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}). \end{aligned} \tag{1.3.52}$$

再使用上面同样的证法,可证

$$P(\sqrt{n}\sigma_0^{-1}|U_{n2i}(t|s)| > \alpha(n\log n)^{-\frac{1}{2}}) = o(n^{-\frac{1}{2}}), \tag{1.3.53}$$

此处,我们省略其证明细节.

综合(1.3.43),(1.3.46),(1.3.48)-(1.1.53),引理1.3.2得证.

# 第2章 乘积限估计的泛函

## §2.1 概率密度核估计的矩与概率不等式及其光滑 Bootstrap 逼近

### 2.1.1 记号与定义

设 $X_1, X_2, \cdots, X_n$ 是非负独立同分布表示寿命的随机变量,其分布函数为 $F$, $Y_1, Y_2, \cdots, Y_n$ 是非负独立同分布表示删失的随机变量,具有连续分布函数 $G$.本文假定诸 $X_i$ 独立于诸 $Y_i$,$F$ 有概率密度 $f$. 在随机右删失模型中,我们不能完全观察 $X_i$,而仅能观察到

$$Z_i = \min(X_i, Y_i), \quad \delta_i = I[X_i \leqslant Y_i], \quad i = 1,2,\cdots,n,$$

$I[\cdot]$ 表示某事件的示性函数,易知 $Z_i$ 的分布函数为 $H = 1-(1-F)(1-G)$. 基于 $(Z_i, \delta_i), i = 1,2,\cdots,n$, Blum 与 Susarla[7] 定义了 $f(t)$ 的如下估计

$$\hat{f}_n(t) = h_n^{-1}\int_0^{\infty} K\left(\frac{t-s}{h_n}\right)\mathrm{d}\hat{F}_n(s), \qquad (2.1.1)$$

其中 $h_n$ 是趋于零的常数序列，$K(\cdot)$ 是核函数，$\hat{F}_n$ 是(0.1) 所定义的 Kaplan-Meier 估计.

我们约定对任何分布函数 $V(\cdot)$，定义 $\overline{V}(\cdot) = 1 - V(\cdot)$，$V^{-r} = (V(\cdot))^{-r}, r > 0, \tau_V = \inf\{t : \overline{V}(t) = 0\}$ 且规定本节中的 $\alpha_0$ 可表示任何所需的绝对常数，$\alpha$ 可表示任何所需的可能与 $F, G$ 有关的常数，即使在同一式中出现也可表示不同. 此外，再定义

$$H(t) = P(Z \leqslant t),$$

$$H_1(t) = P(Z \leqslant t, \delta = 1),$$

$$H_{n1}(t) = \frac{1}{n}\sum_{i=1}^{n} I[Z_i \leqslant t, \delta_i = 1],$$

$$H_0(t) = P(Z \leqslant t, \delta = 0),$$

$$H_{n0} = \frac{1}{n}\sum_{i=1}^{n} I[Z_i \leqslant t, \delta_i = 0].$$

### 2.1.2 $L_1$ 矩与 $p$ 阶绝对矩不等式

**引理 2.1.1** 设 $W$ 是二项 $b(k,p)$ 随机变量，则对 $r \geqslant 1$，有

$$E(1 + W)^{-r} \leqslant r!(kp)^{-r}.$$

证明：见 Susarla 与 Ryzin[75].

**定理 2.1.1** 设 $f(t)$ 存在 $k$ 阶导数，且 $f^{(k)}(t)$ 有界，若核函数满足

i) $K(u)$ 具有有界的支撑$[-1,1]$，且有界；

ii) $\int_{-1}^{1} K(u)\mathrm{d}u = 1$；

iii) 若 $k > 1, \int_{-1}^{1} u^l K(u)\mathrm{d}u = 0, 1 \leqslant l \leqslant k-1, \int_{-1}^{1} |u|^k K(u)\mathrm{d}u = 1$，则对任何 $0 < \tau < \tau_H$，若取 $h_n = n^{-1/[2(k+1)]}$，当 $n$ 充分大时

$$E\int_0^{\tau} |\hat{f}_n(t) - f(t)|\mathrm{d}t$$

$$\leqslant \alpha_0(\overline{G}^{-1}(\tau)\overline{H}^{-3}(\tau)\sup_u|K(u)| + \tau\overline{G}^{-1}(\tau)\sup_u|K(u)|)$$

$$+ \tau \sup_t |f^{(k)}(t)| \left(\int_{-1}^{1} |u^k K(u)| \mathrm{d}u\right) n^{-\frac{k}{2(k+1)}}$$

$$+ O(n^{-\frac{2k+1}{2(k+1)}}).$$

证明:由于 $\hat{F}_n(t)$ 于 $Z_i$ 处的跳为 $\delta_i\{n\bar{\hat{G}}_n(Z_i)\}^{-1}$,此处 $\hat{G}_n$ 是以 $1-\delta_i$ 取代 $\hat{F}_n$ 中 $\delta_i$ 而得到,$\bar{\hat{G}}_n(z_i) = 1 - \hat{G}_n(Z_i)$. 于是 $\hat{f}_n(t)$ 等价地表示为

$$\hat{f}_n(t) = \frac{1}{nh_n}\sum_{i=1}^{n} \frac{\delta_i K\left(\frac{t - Z_i}{h_n}\right)}{\bar{\hat{G}}_n(Z_i)}, \tag{2.1.2}$$

由此等价式,我们可得

$$E\int_0^{\tau} |\hat{f}_n(t) - f(t)| \mathrm{d}t$$

$$\leqslant (nh_n)^{-1}\sum_{i=1}^{n} E\left[ |\bar{\hat{G}}_n^{-1}(Z_i) - \bar{G}^{-1}(Z_i)| \delta_i \int_0^{\tau} K\left(\frac{t - Z_i}{h_n}\right) \mathrm{d}t\right]$$

$$+ E\int_0^{\tau_H} \left| h_n^{-1} \int \frac{K\left(\frac{t-y}{h_n}\right)}{\bar{G}(y)} \mathrm{d}(H_{n1}(y) - H_1(y))\right| \mathrm{d}t$$

$$+ \int_0^{\tau_H} | h_n^{-1} \int K\left(\frac{t-y}{h_n}\right) f(y)\mathrm{d}y - f(t)| \mathrm{d}t \stackrel{\mathrm{d}}{=} \Delta_{n1} + \Delta_{n2} + \Delta_{n3}. \tag{2.1.3}$$

由于 $\bar{\hat{G}}_n^{-1}(Z_i) \leqslant \frac{n+1}{N^+(Z_i)+1}$,于是我们有

$$|\bar{\hat{G}}_n^{-1}(Z_i) - \bar{G}^{-1}(Z_i)|$$

$$\leqslant \frac{n+1}{\bar{G}(Z_i)(N^+(Z_i)+1)} |\hat{G}_n(Z_i) - G(Z_i)|, \tag{2.1.4}$$

注意到随机右删失模型是左截断右删失模型在左截断不发生时的特殊情形,并注意到 $\hat{F}_n$ 与 $\hat{G}_n$ 在理论研究上的对称性,因而由 Gijbels 与 Wang[37] 知

$$\hat{G}_n(t) - G(t) = \bar{G}(t)\frac{1}{n}\sum_{j=1}^{n}\zeta(Z_j,\delta_j;t) + R_n(t), \tag{2.1.5}$$

其中

$$\zeta(Z_j,\delta_j;t) = I[Z_j \leqslant t, \delta_j = 0]/\bar{H}(t) + \int_0^{Z_j \wedge t}\frac{1}{\bar{H}^2(u)}\mathrm{d}H_0(u), \tag{2.1.6}$$

而 $R_n$ 满足

$$E(\sup_{0\leqslant t\leqslant \tau_0}|R_n(t)|^r) = O(n^{-r}), r > 0, 0 < \tau_0 < \tau_H, \tag{2.1.7}$$

因而由(2.1.5)—(2.1.7),并注意 $\zeta(Z_j,\delta_j;t), j = 1,2,\cdots,n$ 是独立同分布且均值为零,方差为 $\int_0^t \frac{\mathrm{d}H_0(s)}{\bar{H}^2(s)}$ 的随机变量,可得

$$\begin{aligned}
&E[(\hat{G}_n(Z_i) - G(Z_i))^2 | Z_i, \delta_i]\\
&\leqslant 2\left\{\frac{\bar{G}^2(Z_i)}{n^2}\left[(n-1)\int_0^{Z_i}\frac{\mathrm{d}H_0(s)}{\bar{H}^2(s)}\right.\right.\\
&\quad + \frac{I[\delta_i = 0]}{\bar{H}(t)}\left(\frac{1}{\bar{H}(t)} - 2\int_0^{Z_i}\frac{1}{\bar{H}^2(t)}\mathrm{d}H_0(s)\right)\\
&\quad \left.\left. + \left(\int_0^{Z_i}\frac{1}{\bar{H}^2(s)}\mathrm{d}H_0(s)\right)^2\right] + E(R_n^2(Z_i)|Z_i,\delta_i)\right\}.
\end{aligned} \tag{2.1.8}$$

于是由(2.1.4),Schward 不等式及(2.1.8),并注意应用引理 2.1.1,即得

$$\begin{aligned}
&E[|\bar{\hat{G}}_n^{-1}(Z_i) - \bar{G}^{-1}(Z_i)| Z_i, \delta_i]\\
&\leqslant \bar{G}^{-1}(Z_i)E^{\frac{1}{2}}\left[\left(\frac{N^+(Z_i)+1}{n+1}\right)^{-2}\middle| Z_i,\delta_i\right]\\
&\quad \times E^{\frac{1}{2}}\left(|\hat{G}_n(Z_i) - G(Z_i)|^2 \middle| Z_i, \delta_i\right)\\
&\leqslant \sqrt{2}\,\frac{n+1}{n-1}\bar{G}^{-1}(Z_i)\bar{H}^{-1}(Z_i)\left\{\frac{\bar{G}^2(Z_i)}{n^2}\left[n\int_0^{Z_i}\frac{\mathrm{d}H_0(s)}{\bar{H}^4(s)}\right.\right.
\end{aligned}$$

$$+\frac{I[\delta_i=0]}{\overline{H}(Z_i)}\left(\frac{1}{\overline{H}(Z_i)}-2\int_0^{Z_i}\frac{1}{\overline{H}^2(s)}\mathrm{d}H_0(s)\right)\Big]$$

$$+E(R_n^2(Z_i)\mid Z_i,\delta_i)\}^{\frac{1}{2}},\tag{2.1.9}$$

注意到(2.1.3) 中所定义的

$$\Delta_{n1}=\frac{1}{nh_n}\sum_{i=1}^{n}E\left[\delta_i\int_0^{\tau}K\left(\frac{t-Z_i}{h_n}\right)\mathrm{d}tE\left(\left|\hat{\overline{G}}^{-1}(Z_i)-\overline{G}^{-1}(Z_i)\right|\ \middle|Z_i,\delta_i\right)\right],$$

(2.1.9) 代入上式,我们有

$$\Delta_{n1}\leqslant\frac{3\sqrt{2}}{h_n}\int_0^{\tau}\int_{t-h_n}^{t+h_n}K\left(\frac{t-y}{h_n}\right)\overline{G}^{-1}(y)\overline{H}^{-1}(y)$$

$$\times\left[\overline{G}(y)\left(\int_0^y\frac{\mathrm{d}H_0(s)}{\overline{H}^4(s)}\right)^{\frac{1}{2}}n^{-\frac{1}{2}}\right.$$

$$\left.+(ER_n^2(y))^{\frac{1}{2}}\right]\mathrm{d}H_1(y)\mathrm{d}t$$

$$\leqslant\frac{3\sqrt{2}}{h_n}\int_0^{\tau}\overline{G}^{-1}(t+h_n)\overline{H}^{-1}(t+h_n)$$

$$\times\overline{G}(t-h_n)\left(\int_0^{t+h_n}\frac{\mathrm{d}H_0(s)}{\overline{H}^4(s)}\right)^{\frac{1}{2}}n^{-\frac{1}{2}}$$

$$+\Big(\sup_{t-h_n\leqslant y\leqslant t+h_n}ER_n^2(y)\Big)^{\frac{1}{2}}\int_{t-h_n}^{t+h_n}K\left(\frac{t-y}{h_n}\right)\overline{G}(y)f(y)\mathrm{d}y\mathrm{d}t.\tag{2.1.10}$$

取 $\tau_0$ 使 $\tau<\tau_0<\tau_H$ 且 $\overline{G}(\tau_0)>\frac{1}{2}\overline{G}(\tau),\overline{H}(\tau_0)>\frac{1}{2}\overline{H}(\tau)$, $\int_0^{\tau_0}\frac{\mathrm{d}H_0(s)}{\overline{H}^4(s)}\leqslant2\int_0^{\tau}\frac{\mathrm{d}H_0(s)}{\overline{H}^4(s)}$(这由它们的连续性可保证),再利用(2.1.7),(2.1.10) 知 $n$ 充分大(以保证 $\tau+h_n<\tau$) 时

$$\Delta_{n1}\leqslant3\sqrt{2}\,\overline{G}^{-1}(\tau_0)\overline{H}^{-1}(\tau_0)[\overline{H}^{-2}(\tau_0)n^{-\frac{1}{2}}+O(\frac{1}{n})]$$

$$\times \int_0^\tau \int_{-1}^1 K(u)\bar{G}(t-h_n u)f(t-h_n u)\mathrm{d}u\mathrm{d}t$$

$$\leqslant 24\bar{G}^{-1}(\tau)\bar{H}^{-3}(\tau)\sup_u |K(u)| h_n^{-1} n^{-\frac{1}{2}}$$

$$+ O(n^{-1}h_n^{-1}), \tag{2.1.11}$$

又

$$\Delta_{n2} \leqslant h_n^{-1}\int_0^\tau E^{\frac{1}{2}} \left| \int_{t-h_n}^{t+h_n} \frac{K\left(\frac{t-y}{h_n}\right)}{\bar{G}(y)} \mathrm{d}(H_{n1}(y) - H_1(y)) \right|^2 \mathrm{d}t$$

$$\leqslant h_n^{-1}\int_0^\tau [E\int_0^\infty \int_0^\infty I[t-h_n \leqslant y_1 \leqslant t+h_n]$$

$$\times I[t-h_n \leqslant y_2 \leqslant t+h_n]$$

$$\times \frac{K\left(\frac{t-y_1}{h_n}\right)K\left(\frac{t-y_2}{h_n}\right)}{\bar{G}(y_1)\bar{G}(y_2)} \mathrm{d}(H_{n1}(y_1)$$

$$- H_1(y_1))\mathrm{d}(H_{n1}(y_2) - H_1(y_2))]^{\frac{1}{2}}\mathrm{d}t. \tag{2.1.12}$$

令

$$h_{n0}(y_1,y_2) = I[t-h_n \leqslant y_1 \leqslant t+h_n]$$

$$\times I[t-h_n \leqslant y_2 \leqslant t+h_n] \frac{K\left(\frac{t-y}{h_n}\right)K\left(\frac{t-y_2}{h_n}\right)}{\bar{G}(y_1)\bar{G}(y_2)},$$

由于 $G$ 连续,因而在 $n$ 充分大时 $\bar{G}(t+h_n) > \frac{1}{2}\bar{G}(t)$,由此在 $t \leqslant \tau$ 时可推得

$$|h_{n0}(y_1,y_2)| \leqslant 4(\sup_u |K(u)|)^2 \bar{G}^{-2}(\tau).$$

于是 Chang 与 Rao[16] 附录中的引理可用于计算(2.1.12)中的期望值,由此引理我们得

$$\Delta_{n2} \leqslant 2\tau\alpha_0 \sup_u |K(u)| \bar{G}^{-1}(\tau) n^{-\frac{1}{2}} h_n^{-1}. \tag{2.1.13}$$

而

$$|\Delta_{n3}| = |h_n^k\int_0^\tau\int_{-1}^1 u^k K(u) f^{(k)}(t-\theta h_n u)\mathrm{d}u\mathrm{d}t|$$
$$\leqslant [\tau \sup_t |f^{(k)}(t)| \int_{-1}^1 |u^k K(u)|\mathrm{d}u] h_n^k,$$

综合(2.1.3),(2.1.11),(2.1.13),(2.1.14)并取 $h_n = n^{-\frac{1}{2(k+1)}}$,即得定理 2.1.1 的证明.

**定理 2.1.2** 设 $f(t)$ 存在 $k$ 阶导数,且 $f^{(k)}(t)$ 有界,$k\geqslant 1$,若 $K(\cdot)$ 满足定理2.1.1的条件,则对任何取定的 $0<t<\tau_H$,对 $p\geqslant 2$,当 $n$ 充分大时

$$E|\hat{f}_n(t)-f(t)|^p \leqslant \alpha_0(\bar{G}^{1-p}(t)\bar{H}^{-3p}(t)h_n^{-p}n^{-\frac{p}{2}} + h_n^k).$$

**推论** 若定理2.1.2的条件满足,则 $h_n = n^{-\frac{1}{2(k+1)}}$ 时,对取定的 $0<t<\tau_H$ 及 $p\geqslant 2$,有

$$E|\hat{f}_n(t)-f(t)|^p \leqslant \alpha_0\bar{G}^{1-p}(t)\bar{H}^{-3p}(t)n^{-\frac{kp}{2(k+1)}}.$$

定理 2.1.2 的证明:使用 $\hat{f}_n(t)$ 的等价表达式(2.1.2),有

$$\hat{f}_n(t)-f(t)$$
$$=\left[\frac{1}{nh_n}\sum_{i=1}^n \frac{\delta_i K\left(\frac{t-Z_i}{h_n}\right)}{\bar{\hat{G}}_n(Z_i)} - \frac{1}{nh_n}\sum_{i=1}^n \frac{\delta_i K\left(\frac{t-Z_i}{h_n}\right)}{\bar{G}(Z_i)}\right]$$
$$+\left[\frac{1}{nh_n}\sum_{i=1}^n \frac{\delta_i K\left(\frac{t-Z_i}{h_n}\right)}{\bar{G}(Z_i)} - \frac{1}{nh_n}\sum_{i=1}^n E\frac{\delta_i K\left(\frac{t-Z_i}{h_n}\right)}{\bar{G}(Z_i)}\right]$$
$$+\left[\frac{1}{nh_n}\sum_{i=1}^n E\frac{\delta_i K\left(\frac{t-Z_i}{h_n}\right)}{\bar{G}(Z_i)} - f(t)\right]$$
$$\stackrel{\mathrm{d}}{=} E_{1n}+E_{2n}+E_{3n}. \tag{2.1.15}$$

对 $p\geqslant 2$,利用 $C_p$ 不等式,得

$$E|E_{1n}|^p \leqslant h_n^{-p}n^{-1}\sum_{i=1}^n E\{\delta_i K^p\left(\frac{t-Z_i}{h_n}\right)$$

$$\cdot E(|\hat{G}_n^{-1}(Z_i)-\bar{G}^{-1}(Z_i)|^p|Z_i,\delta_i)\},\tag{2.1.16}$$

由王启华[107] 中(3.50) 知

$$E(|\bar{G}_n^{-1}(Z_i)-\bar{G}^{-1}(Z_i)|^p|Z_i,\delta_i)\leqslant\alpha_0 n^{-\frac{p}{2}}\bar{G}^{-p}(Z_i)\bar{H}^{-3p}(Z_i).\tag{2.1.17}$$

将(2.1.17) 代入(2.1.16) 即知 $n$ 充分大时

$$\begin{aligned}E|E_{1n}|^p&\leqslant\alpha_0 h_n^{-p}n^{-\frac{p}{2}}\int_{t-h_n}^{t+h_n}\frac{K^p\left(\frac{t-s}{h_n}\right)}{\bar{G}^{p-1}(s)\bar{H}^{3p}(s)}f(s)\mathrm{d}s\\&\leqslant\alpha_0 h_n^{-p}n^{-\frac{p}{2}}\bar{G}^{1-p}(t+h_n)\bar{H}^{-3p}(t+h_n)\int_{t-h_n}^{t+h_n}f(s)\mathrm{d}s\\&\leqslant\alpha_0 h_n^{-p}n^{-\frac{p}{2}}\bar{G}^{1-p}(t)\bar{H}^{-3p}(t).\end{aligned}\tag{2.1.18}$$

上式利用到 $K(\cdot)$ 的有界性及 $G,H$ 的连续性,因 $G,H$ 的连续性保证了 $\bar{G}(t+h_n)>\frac{1}{2}\bar{G}(t),\bar{H}(t+h_n)>\frac{1}{2}\bar{H}(t)$ 在 $n$ 充分大时成立.以后将会多次利用类似的性质,但不再特别声明.

又应用 Dharmadhikari-Jodgeo(D-J) 不等式,并再次利用 $K$ 的有界性和 $\bar{G}$ 的连续性可得

$$\begin{aligned}E|E_{2n}|^p&\leqslant\alpha_0(nh_n)^{-p}n^{\frac{p}{2}-1}\sum_{i=1}^{n}E\left|\frac{\delta_iK\left(\frac{t-Z_i}{h_n}\right)}{\bar{G}(Z_i)}-E\frac{\delta_iK\left(\frac{t-Z_i}{h_n}\right)}{\bar{G}(Z_i)}\right|^p\\&\leqslant\alpha_0 n^{-\frac{p}{2}}h_n^{-p}\int\frac{K^p\left(\frac{t-y}{h_n}\right)}{\bar{G}^{p-1}(y)}\mathrm{d}F(y)\\&\leqslant\alpha_0 n^{-\frac{p}{2}}h_n^{-p}\bar{G}^{1-p}(t).\end{aligned}\tag{2.1.19}$$

而

$$E_{3n}=h_n^{-1}\int K\left(\frac{t-y}{h_n}\right)\mathrm{d}F(y)-f(t),$$

利用定理中的核条件(i),(ii),(iii) 即得

$$
\begin{aligned}
|E_{3n}| &= \left|\int_{-1}^{1} K(u)(f(t-h_n u)-f(t))\,\mathrm{d}u\right| \\
&= h_n^k \left|\int_{-1}^{1} u^k K(u) f^k(t-\theta h_n u)\mathrm{d}u\right| \\
&\leqslant \alpha_0 h_n^k. \qquad (2.1.20)
\end{aligned}
$$

于是综合(2.1.15),(2.1.18),(2.1.19) 与(2.1.20) 并运用 $C_p$ 不等式,即得定理 2.1.2 的结论.

推论的证明:只要在定理 2.1.2 中取 $h_n = n^{-\frac{1}{2(k+1)}}$,即可得推论 2.1.1 的证明. 又注意到 $\overline{G}^{1-p}(t)\overline{H}^{-3p}(t) \geqslant 1$,从而将两项合为一项(这在证上面定理 2.1.2 时也用到类似的手法).

### 2.1.2 $\hat{f}_n(t)$ 的偏差概率不等式

**定理 2.1.3** 设 $f(t)$ 连续,若 $K(\cdot)$ 是具有有界支撑$[-1,1]$的概率密度核函数,则对 $\forall\, \varepsilon > 0$,及任取定的 $0 < t < \tau_H$,知 $n$ 充分大时

$$
\begin{aligned}
& P(|\hat{f}_n(t) - f(t)| > \varepsilon) \\
\leqslant & \alpha_0 \exp\{-\alpha_0 \overline{G}^2(t)\overline{H}^2(t)[\min(1,\varepsilon)]^2 n h_n^2\}.
\end{aligned}
$$

为证该定理,我们先证下面的

**引理 2.1.2** 设对某 $\gamma > 0$,若 $\tau_0$ 满足 $\overline{H}(\tau_0) > \gamma$,则当$\frac{\sqrt{50}}{n\gamma} < \eta < \sqrt{2}$ 时

$$
P(\sup_{0\leqslant y\leqslant \tau_0} |\hat{\overline{G}}_n(y) - \overline{G}(y)| > \eta) \leqslant \alpha_0 \exp(-\alpha_0 \gamma^2 \eta^2 n).
$$

证明:利用不等式 $|x-y| \leqslant |\log x - \log y|$,$0 < x < 1$,$0 < y < 1$,我们有

$$
\begin{aligned}
& P(\sup_{0\leqslant y\leqslant \tau_0} |\hat{\overline{G}}_n(y) - \overline{G}(y)| > \eta) \\
\leqslant & P(\sup_{0\leqslant y\leqslant \tau_0} |\log\hat{\overline{G}}_n(y) - \log\overline{G}(y)| > \eta)
\end{aligned}
$$

$$\cdot P(\sup_{0\leqslant y\leqslant \tau_0} n\,|\log\hat{\bar{G}}_n(y) - \log\bar{G}(y)|^2 > n\eta^2). \quad (2.1.21)$$

而由 Major 与 Rejtö[54] 中(2.18) 即得

$$P(\sup_{0\leqslant y\leqslant \tau_0} n\,|\log\hat{\bar{G}}_n(y) - \log\bar{G}(y)|^2 > n\eta^2)$$

$$\leqslant \alpha_0 \exp\{-\alpha_0\gamma^2 x^2 n\}, \quad (2.1.22)$$

于是(2.1.21),(2.1.22)一起就证明了引理 2.1.2.

定理 2.1.3 的证明:利用(2.1.15),由于对任取但固定的 $0 < t < \tau_H$,

$$E_{3n} = \int_{-1}^{1} K(u)(f(t-h_n u) - f(t))\mathrm{d}u, \quad (2.1.23)$$

而由 $f(t)$ 的连续性知,对 $\forall \varepsilon > 0$,当 $n$ 充分大时

$$|f(t-h_n u) - f(t)| < \frac{\varepsilon}{3},$$

对 $-1 \leqslant u \leqslant 1$ 一致地成立,由此及(2.1.23) 即知

$$E_{3n} \leqslant \frac{\varepsilon}{3}, \quad (2.1.24)$$

于是由(2.1.15),(2.1.24) 知对上面 $\varepsilon > 0$

$$P(|\hat{f}_n(t) - f(t)| > \varepsilon)$$

$$\leqslant P\left(|E_{1n}| > \frac{\varepsilon}{3}\right) + P\left(|E_{2n}| > \frac{\varepsilon}{3}\right), \quad (2.1.25)$$

由于

$$E_{1n} \leqslant \sup_{t-h_n u\leqslant y\leqslant t+h_n u} |\hat{\bar{G}}_n(y) - \bar{G}(y)|\,\hat{\bar{G}}_n^{-1}(t+h_n)\frac{1}{nh_n}$$

$$\cdot \sum_{i=1}^{n} \frac{\delta_i K\left(\frac{t-Z_i}{h_n}\right)}{\bar{G}(Z_i)},$$

取 $\tau_0$ 使 $t < \tau_0 < \tau_H$ 且 $\bar{G}(\tau_0) \geqslant \frac{1}{2}\bar{G}(t)$,$\bar{H}(\tau_0) \geqslant \frac{1}{2}\bar{H}(t)$,则 $n$ 充分大(以保证 $t + h_n < \tau_0$) 时,有

$$P(|E_{1n}| > \frac{\varepsilon}{3})$$

$$\leqslant P(\sup_{0\leqslant y\leqslant \tau_0} |\hat{\bar{G}}_n(y)-\bar{G}(y)|\hat{\bar{G}}_n^{-1}(\tau_0)>\frac{1}{6})$$

$$+P\left(\left|\frac{1}{nh_n}\sum_{i=1}^{n}\frac{\delta_i K\left(\frac{t-Z_i}{h_n}\right)}{\bar{G}(Z_i)}-h_n^{-1}\int K\left(\frac{t-y}{h_n}\right)f(y)\mathrm{d}y\right|>\varepsilon\right)$$

$$+P\left(\sup_{0\leqslant y\leqslant \tau_0}|\hat{\bar{G}}_n(y)-\bar{G}(y)|h_n^{-1}\int K\left(\frac{t-y}{h_n}\right)f(y)\mathrm{d}y>\frac{\varepsilon}{6}\right)$$

$$\stackrel{\mathrm{d}}{=} I_{1n}+I_{2n}+I_{3n}. \tag{2.1.26}$$

由于 $\bar{G}(\tau_0)>\frac{1}{2}\bar{G}(t)$,因而

$$I_{1n}\leqslant P\left(\sup_{0\leqslant y\leqslant \tau_0}|\hat{\bar{G}}_n(y)-\bar{G}(y)|>\frac{\bar{G}(t)}{24}\right)$$

$$+P(\hat{\bar{G}}_n(\tau_0)\leqslant\frac{1}{2}\bar{G}(\tau_0)). \tag{2.1.27}$$

又

$$P(\hat{\bar{G}}_n(\tau_0)\leqslant\frac{1}{2}\bar{G}(\tau_0))$$

$$\leqslant P(\hat{\bar{G}}_n(\tau_0)-\bar{G}(\tau_0))|>\frac{1}{2}\bar{G}(\tau_0))$$

$$\leqslant P\left(\sup_{0\leqslant y\leqslant \tau_0}|\hat{\bar{G}}_n(y)-\bar{G}(y)|>\frac{\bar{G}(t)}{4}\right). \tag{2.1.28}$$

在引理 2.1.2 中取 $\gamma=\frac{1}{2}\bar{H}(\tau_0)$,则由(2.1.27),(2.1.28)并利用引理 2.1.2,知 $n$ 充分大(以保证$\frac{\sqrt{50}}{n\gamma}<\frac{1}{4}\bar{G}(t)$)时

$$I_{1n}\leqslant \alpha_0\exp\{-\alpha_0(\frac{1}{2}\bar{H}(\tau_0))^2\bar{G}^2(t)n\}$$

$$\leqslant \alpha_0\exp\{-\alpha_0\bar{H}^2(t)\bar{G}^2(t)n\}. \tag{2.1.29}$$

记 $A_n=\{y:H_n(y)-H(y^-)>0\}$,$A_n^c$ 为其补,则 $n$ 充分大时

$$I_{2n}\leqslant P\left(\left|\frac{1}{h_n}\int\frac{K\left(\frac{t-y}{h_n}\right)}{\bar{G}(y)}\mathrm{d}(H_{n1}(y)-H_1(y))\right|>\varepsilon\right)$$

$$\leqslant P\left\{\left|\frac{1}{h_n}\int_{A_n}\frac{K\left(\frac{t-y}{h_n}\right)}{\bar{G}(y)}\mathrm{d}(H_{n1}(y)-H_1(y))\right|>\frac{h_n\varepsilon}{2}\right\}$$

$$+P\left\{\left|\frac{1}{h_n}\int_{A_n^c}\frac{K\left(\frac{t-y}{h_n}\right)}{\bar{G}(y)}\mathrm{d}(H_{n1}(y)-H_1(y))\right|>\frac{h_n\varepsilon}{2}\right\}$$

$$\leqslant 2P(\bar{G}^{-1}(t+h_n)\sup_u K(u)\cdot\sup_y|H_{n1}(y)-H_1(y)|>\frac{h_n\varepsilon}{2})$$

$$\leqslant 2\exp\left\{-2n\frac{\bar{G}^2(t+h_n)}{4(\sup_u K(u))^2}h_n^2\varepsilon^2\right\}$$

$$\leqslant 2\exp\left\{-\frac{\bar{G}^2(t)\varepsilon^2}{8(\sup_u K(u))^2}nh_n^2\right\}. \qquad (2.1.30)$$

上面最后一个不等式应用到 $\bar{G}(t+h_n)\geqslant\frac{1}{2}\bar{G}(t)$ 在 $n$ 充分大时成立这一事实. 倒数第二个不等式应用到 Dvoretzky, Kiefer 与 Wolfowitz 的关于经验分布函数的概率不等式.

由(2.1.24)知

$$\frac{1}{h_n}\int K\left(\frac{t-y}{h_n}\right)f(y)\,\mathrm{d}y\to f(t),n\to\infty,$$

因而 $n$ 充分大时

$$\frac{1}{h_n}\int K\left(\frac{t-y}{h_n}\right)f(y)\,\mathrm{d}y\leqslant 2f(t),$$

故由(2.1.26)定义的

$$I_{3n}\leqslant P\left(\sup_{0\leqslant y\leqslant\tau_0}|\hat{\bar{G}}(y)-\bar{G}(y)|>\frac{\varepsilon}{12f(t)}\right),$$

再次应用引理 2.1.2 并取 $\gamma=\frac{1}{2}\bar{H}(\tau_0)>\frac{1}{2}\bar{H}(t)$, 则 $n$ 充分大 (以保证 $\frac{\sqrt{50}}{n\gamma}<\min\left\{1,\frac{\varepsilon}{12f(t)}\right\}$) 时

$$I_3\leqslant\alpha_0\exp\left\{-\alpha_0\bar{H}^2(t)\left(\min\left\{1,\frac{\varepsilon}{12f(t)}\right\}\right)^2n\right\}. \qquad (2.1.31)$$

综合(2.1.26),(2.1.29),(2.1.31)与(2.1.32),知 $n$ 充分大(以保证对所固定的 $0 < t < \tau_H$,有 $h_n f(t) \leqslant 1$)时,有

$$P(|E_{1n}| > \frac{\varepsilon}{3}) \leqslant \alpha_0 \exp\{-\alpha_0 \overline{G}^2(t)\overline{H}^2(t)[\min(1,\varepsilon)]^2 nh_n^2\}, \tag{2.1.32}$$

而由(2.1.30)的证明,即知

$$P(|E_{2n}| > \frac{\varepsilon}{3}) \leqslant 2\exp\left\{-\frac{\overline{G}^2(t)\varepsilon^2}{72\sup\limits_u K(u)} nh_n^2\right\}, \tag{2.1.33}$$

于是综合(2.1.25),(2.1.23)与(2.1.24),得

$$\begin{aligned} & P(|\hat{f}_n(t) - f(t)| > \varepsilon) \\ \leqslant & \alpha_0 \exp\{-\alpha_0 \overline{G}^2(t)\overline{H}^2(t)[\min(1,\varepsilon)]^2 nh_n^2\}. \end{aligned}$$

这就证明了定理2.1.3.

注:在上面定理证明过程中,每次取 $n$ 充分大时,都认为比前一次取得大.

### 2.1.3 光滑 Bootstrap 逼近

本节探讨 $\hat{f}_n(t)$ 的光滑 bootstrap 逼近问题. 众知,在随机删失模型下引进 bootstrap 方法已有一些工作(见 Efron[28],Lo 与 Singh[51]),这种一般的 bootstrap 方法在某些场合并不适用,如在需要 bootstrap 样本分布具有某种光滑性时就是如此. 在此我们再次将 Efron[27] 在完全样本情况下所提出的光滑 bootstrap 思想应用到我们的删失模型,证明了 $\hat{f}_n(t)$ 的光滑 bootstrap 逼近成立,且证明了概率密度 bootstrap 估计的方差几乎处处收敛到 $\hat{f}_n(t)$ 的渐近方差. 显然本节的结果可用于寻求 $f(t)$ 的区间估计及 $f(t)$ 的假设检验等问题,因而是很有意义的.

记 $\hat{G}_n$ 为 $G$ 的乘积限估计,它是以 $1-\delta_i$ 取代 $\hat{F}_n$ 中的 $\delta_i$ 而得到. 且记

$$\widehat{SF}_n = \int_0^t \hat{f}_n(y)\mathrm{d}y,$$

$$\hat{g}_n(t) = h_n^{-1}\int_0^\infty K\left(\frac{t-y}{h_n}\right)\mathrm{d}\hat{G}_n(y),$$

$$\widehat{SG}_n(t) = \int_0^t g_n(y)\mathrm{d}y.$$

为抽取第二阶段 bootrstrap 样本，我们将 $\widehat{SF}_n$，$\widehat{SG}_n$ 分别看作真的生存和删失总体，根据随机右删失模型我们得到独立同分布的 bootrstrap 观察 $(Z_i^*,\delta_i^*),\cdots,(Z_n^*,\delta_n^*)$，此处

$$Z_i^* = \min(X_i^*,Y_i^*),\quad \delta_i^* = I[X_i^* \leqslant Y_i^*], i = 1,2,\cdots,m.$$

$m$ 可以不等于 $n$（事实上在此是独立于 $n$ 取值）. 而 $X_1^*,X_2^*,\cdots,X_m^* \overset{\text{iid}}{\sim} \widehat{SF}_n$，$Y_1^*,Y_2^*,\cdots,Y_m^* \overset{\text{iid}}{\sim} \widehat{SG}_n$，且诸 $X_i^*$ 独立于诸 $Y_i^*$. 以下均以 $P^*,E^*,\mathrm{Var}^*$ 表示在给定 $(Z_1,\delta_1),\cdots,(Z_n,\delta_n)$ 的条件下，在 bootrstrap 概率空间中求概率，期望和方差，而 $\mathcal{L}^*$ 表示 bootrstrap 分布弱收敛，$a.s.^*$ 表示以 $P^*$ 概率 1. 此外，再记

$$\widetilde{SH}_n = P^*(Z_i^* \leqslant t), \widehat{SH}_{nm}^* = \frac{1}{m}\sum_{i=1}^m I[Z_i^* \leqslant t],$$

$$\widetilde{SH}_{n0} = P^*(Z_i^* \leqslant t, \delta_i^* = 0),$$

$$\widehat{SH}_{nm0}^* = \frac{1}{m}\sum_{i=1}^m I[Z_i^* \leqslant t, \delta_i^* = 0],$$

$$\widetilde{SH}_{n1} = P^*(Z_i^* \leqslant t, \delta_i^* = 1),$$

$$\widehat{SH}_{nm1}^* = \frac{1}{m}\sum_{i=1}^m I[Z_i^* \leqslant t, \delta_i^* = 1].$$

$\hat{F}_{nm}^*$，$G_{nm}^*$ 是基于 $(Z_i^*,\delta_i^*)$，$i = 1,2.\cdots,m$ 的 bootrstrap 乘积限估计，即分别在 $\hat{F}_n$，$\hat{G}_n$ 的表达式中将 $n$ 换成 $m$，并以 $(Z_i^*,\delta_i^*)$ 取代 $(Z_i,\delta_i)$ 就得到 $\hat{F}_{nm}^*$，$G_{nm}^*$.

**引理 2.1.3** 设 $K(\cdot)$ 是具有支撑集 $[-1,1]$ 的连续概率密度函数，则当 $m \to \infty$ 时

(i) $\int_{-1}^1 K(u)\hat{f}_n(t - h_m u)\mathrm{d}u \to \hat{f}_n(t)$;

(ii) $\int_{-1}^{1} K^{2+2\alpha}(u) \dfrac{\hat{f}_n(t-h_m u)}{\widehat{SG}_n^{1+2\alpha}(t-h_m u)} du$

$$\to \frac{\hat{f}_n(t)}{\widehat{SG}_n^{1+2\alpha}(t)} \int_{-1}^{1} K^{2+2\alpha}(u)\, du, \quad \alpha \geqslant 0.$$

证明:证明极其简单,只要主要注意到 $m \to \infty$ 时

$$\hat{f}_n(t-h_m u) \to \hat{f}_n(t), \quad \frac{\hat{f}_n(t-h_m u)}{\widehat{SG}_n^{1+2\alpha}(t-h_m u)} \to \frac{\hat{f}_n(t)}{\widehat{SG}_n^{1+2\alpha}(t)}$$

关于 $-1 \leqslant u \leqslant 1$ 一致成立即可得证.

**引理 2.1.4** (i) 设 $f(t)$ 连续,若 $K(\cdot)$ 是具有有界支撑集 $[-1,1]$,在 $(-1,1)$ 上存在一阶有界导函数的有界概率密度核函数,则对任何取定的 $0 < t < \tau_H$,当 $\dfrac{nh_n^2}{\log\log n} \to \infty$ 时

$$\hat{f}_n(t) \xrightarrow{a.s.} f(t).$$

(ii) 设 $K(\cdot)$ 是具有有界支撑集 $[-1,1]$ 的概率核函数,则对任何取定的 $0 < t < \tau_H$,有

$$\widehat{SG}_n(t) \xrightarrow{a.s.} G(t).$$

注 1:引理 2.1.4(i) 是关于 $\hat{f}_n(t)$ 强相合性的一个结果,这在文献中已有很多研究,但已有的强相合性定理所使用的条件不符合以下诸定理的要求,故不能直接引用,但考虑到证法上的类似性,故引理 2.1.4(i) 的证明从略.

引理 2.1.4(ii) 的证明:由于

$$\begin{aligned} \widehat{SG}_n(t) &= h_n^{-1} \int_0^t \int_{s-h_n}^{s+h_n} K\left(\frac{s-y}{h_n}\right) d\hat{G}_n(y) ds \\ &= \int_0^{t+h_n} h_n^{-1} \int_{y-h_n}^{y+h_n} K\left(\frac{s-y}{h_n}\right) ds d\hat{G}_n(y) \\ &= \hat{G}_n(t+h_n), \end{aligned} \tag{2.1.34}$$

于是

$$|\widehat{SG}_n(t) - G(t)| \leqslant |\hat{G}_n(t+h_n) - G(t+h_n)|$$

$$+|G(t+h_n)-G(t)|,$$

因而由 Csörgo 与 Horváth[23] 中定理 1 及 $G$ 的连续性即得引理 2.1.4(ii) 的证明.

现设 $\hat{f}^*_{nm}(t)$ 是对应于 $\hat{f}_n(t)$ 的 Bootstrap 概率密度,即

$$\hat{f}^*_{nm}(t)=h_n^{-1}\int_0^{\infty}K\left(\frac{t-y}{h_n}\right)\mathrm{d}\hat{F}^*_{nm}(y). \tag{2.1.35}$$

**定理 2.1.4** (a) 设 $K(\cdot)$ 具有有界支撑集$[-1,1]$的有界概率密度核函数,则对任何 $0<t<\tau_H$,当 $mh_m\to\infty$, $mh_m{}^3\to 0$ 时

$$\sqrt{mh_m}(\hat{f}^*_{nm}(t)-\hat{f}_n(t))\xrightarrow{\mathscr{L}^*}N\left(0,\frac{\hat{f}_n(t)}{\widehat{\overline{SG}}_n(t)}\int_{-1}^{1}K^2(u)\mathrm{d}u\right).$$

(b) 设 $f(t)$ 连续,若 $K(\cdot)$ 满足(a) 中的条件,则当 $mh_m^3\to 0$, $\dfrac{nh_n^2}{\log\log n}\to\infty$ 时,对任何取定的 $0<t<\tau_H$,以概率 1,有

$$\sqrt{mh_m}(\hat{f}^*_{nm}(t)-\hat{f}_n(t))\xrightarrow{\mathscr{L}^*}N\left(0,\frac{f(t)}{\overline{G}(t)}\int_{-1}^{1}K^2(u)\mathrm{d}u\right).$$

(c) 若 $f(t)$ 存在一阶有界导数,$K(\cdot)$ 是$[-1,1]$上的有界概率密度核函数,则对任何取定的 $0<t<\tau_H$,当 $nh_n\to\infty$, $nh_n^3\to 0$ 时,有

$$\sqrt{nh_n}(\hat{f}_n(t)-f(t))\xrightarrow{\mathscr{L}}N\left(0,\frac{f(t)}{\overline{G}(t)}\int_{-1}^{1}K^2(u)\mathrm{d}u\right).$$

(d) 若 $f(t)$ 满足(c) 中的条件,$K(\cdot)$ 满足(b) 中的条件,则对任何 $0<t<\tau_H$,当 $mh_m^3\to 0$, $\dfrac{nh_n^2}{\log\log n}\to\infty$ 时,以概率 1 有

$$\sup_x|P^*(\sqrt{mh_m}(\hat{f}^*_{nm}(t)-\hat{f}_n(t))\leqslant x)$$
$$-P(\sqrt{nh_n}(\hat{f}_n(t)-f(t))\leqslant x)|\to 0.$$

注 2:(a) 中假定 $mh_m\to\infty$,而(b) 却没列出此条件,这是因为(b) 现有的条件蕴涵 $mh_m\to\infty$,事实上由 $mh_m^3\to 0$ 知 $m\to\infty$,若此时 $mh_m\not\to\infty$,则 $n\to\infty$ 时 $nh_n\not\to\infty$,于是更有 $\dfrac{nh_n^2}{\log\log n}\not\to\infty$,这

与假设矛盾. (d) 与(c) 比较也有类似的情况.

定理 2.1.4 的证明:(a) 注意到

$$
\begin{aligned}
&\hat{f}^*_{nm}(t) - \hat{f}_n(t) \\
&= \left[\frac{1}{mh_m}\sum_{i=1}^{m}\frac{\delta_i^* K\left(\frac{t-z_i^*}{h_m}\right)}{\overline{\widehat{SG}}_n(Z_i^*)} - \frac{1}{h_m}\int_0^\infty K\left(\frac{t-y}{h_m}\right)\mathrm{d}\,\widehat{SF}_n(y)\right] \\
&\quad + \frac{1}{mh_m}\sum_{i=1}^{m}\delta_i^* K\left(\frac{t-z_i^*}{h_m}\right)\left[\frac{1}{\overline{\hat{G}}^*_m(Z_i^*)} - \frac{1}{\overline{\widehat{SG}}_n(Z_i^*)}\right] \\
&\quad + \left(\frac{1}{h_m}\int_0^\infty K\left(\frac{t-y}{h_m}\right)\mathrm{d}\,\widehat{SF}_n(y) - \frac{1}{h_n}\int_0^\infty K\left(\frac{t-y}{h_n}\right)\mathrm{d}\hat{F}_n(y)\right) \\
&\overset{\mathrm{d}}{=} R^*_{1nm} + R^*_{2nm} + R_{3nm},
\end{aligned}
\tag{2.1.36}
$$

令

$$
U^*_{nmj} = \frac{\delta_i^* K\left(\frac{t-z_i^*}{h_m}\right)}{h_m\overline{\widehat{SG}}_n(Z_i^*)}, W^*_{nmj} = \frac{U^*_{nmj} - EU^*_{nmj}}{\sqrt{m}\,\mathrm{Var}^*\,U^*_{nmj}}, j = 1,2,\cdots,n,
$$

由于

$$
\begin{aligned}
&\mathrm{Var}^*\,U^*_{nmj} \\
&= h_m^{-2}\int_0^\infty \frac{K^2\left(\frac{t-y}{h_m}\right)}{\overline{\widehat{SG}}_n(y)}\mathrm{d}\,\widehat{SF}_n(y) - \frac{1}{h_m^2}\left(\int_0^\infty K\left(\frac{t-y}{h_m}\right)\mathrm{d}\,\widehat{SF}_n(y)\right)^2 \\
&\geqslant h_m^{-2}\int_{t-h_m}^{t+h_m} K^2\left(\frac{t-y}{h_m}\right)\left[\frac{1}{\overline{\widehat{SG}}_n(y)} - 1\right]\mathrm{d}\,\widehat{SF}_n(y) \\
&\geqslant h_m^{-2}\left[\frac{1}{\overline{\widehat{SG}}_n(t-h_m)} - 1\right]\int_0^\infty K^2\left(\frac{t-y}{h_m}\right)\mathrm{d}\,\widehat{SF}_n(y).
\end{aligned}
\tag{2.1.37}
$$

注意到 $Z_j^* > t + h_m$ 时, $U^*_{nmj} = 0$,因而

$|W_{nmj}{}^*|$

$$\leqslant \frac{[\delta_j^* K\left(\frac{t-Z_j^*}{h_m}\right) I[Z_j^* \leqslant t+h_m] + \overline{\widehat{SG}}_n(t+h_m)\int_0^\infty K\left(\frac{t-y}{h_m}\right) \mathrm{d}\,\widehat{SF}_n(y)]\,\overline{\widehat{SG}}^{\frac{1}{2}}(t-h_m)}{\overline{\widehat{SG}}_n(t+h_m)(m\,\widehat{SG}_n(t-h_m)\int_0^\infty K^2\left(\frac{t-y}{h_m}\right)\mathrm{d}\,\widehat{SF}_n(y))^{\frac{1}{2}}}$$

$$\leqslant \frac{2\sup\limits_u K(u)}{\sqrt{mh_m}\,\overline{\widehat{SG}}_n(t+h_m)(\widehat{SG}_n(t-h_m)\int_{-1}^1 K^2(u) f_n(t-h_m u)\mathrm{d}u)^{\frac{1}{2}}} \to 0, \qquad (2.1.38)$$

上面应用到引理2.1.4(ii) 及$\widehat{SG}_n$的连续性. 由(2.1.38)知对任何$\tau > 0$,若定义

$$W_{nmj}^{*\,(\tau)} = \begin{cases} W_{nmj}^*, & 若\,|W_{nmj}^*| \leqslant \tau, \\ 0, & 其他, \end{cases}$$

则当 $m$ 充分大时$W_{nmj}^{*\,(\tau)} = W_{nmj}^*$,于是 $m$ 充分大时

$$\sum_{j=1}^m \mathrm{Var}^*\ W_{nmj}^{*\,(\tau)} = 1, \qquad (2.1.39)$$

$$\sum_{j=1}^m E^*\ W_{nmj}^{*\,(\tau)} = 0. \qquad (2.1.40)$$

又对 $\alpha > 0$

$$\sum_{j=1}^m P^*(|W_{nmj}^*| > \varepsilon) \leqslant \frac{mE^*|U_{nm1}^* - E^*U_{nm1}^*|^{2+2\alpha}}{\varepsilon^{2+2\alpha} m^{1+\alpha} \mathrm{Var}^{*\frac{2+2\alpha}{2}}(U_{nm1}^*)}, \qquad (2.1.41)$$

又

$$\begin{aligned} & E^*|U_{nm1}^*|^{2+2\alpha} \\ &= h_m^{-2(1+\alpha)}\int_{t-h_m}^{t+h_m} \frac{K^{2+2\alpha}\left(\frac{t-y}{h_m}\right)}{\overline{\widehat{SG}}_n^{1+2\alpha}(y)} \hat{f}_n(y)\mathrm{d}y \\ &= h_m^{-1-2\alpha}\int_{-1}^1 K^{2+2\alpha}(u)\frac{\hat{f}_n(t-h_m u)}{\overline{\widehat{SG}}_n^{1+2\alpha}(t-h_m u)}\mathrm{d}u, \qquad (2.1.42) \end{aligned}$$

和

$$\mathrm{Var}^*\,U_{nm1}^* = h_m^{-1}\int_{-1}^1 K^2(u)\frac{\hat{f}_n(t-h_m u)}{\overline{\widehat{SG}}_n(t-h_m u)}\mathrm{d}u$$

$$
-\left(\int_{-1}^{1} K(u)\hat{f}_n(t-h_m u)\mathrm{d}u\right)^2, \qquad (2.1.43)
$$

于是由(2.1.41)—(2.1.43) 及引理 2.1.4 知 $mh_m \to \infty$ 时

$$
\sum_{j=1}^{m} P^*(|W_{nmj}^*| > \varepsilon) \leqslant \frac{2^{1+2a}h_m^{1+2a}(E^*|U_{nm1}^*|^{2+2a} + |E^* U_{nm1}^*|^{2+2a})}{\varepsilon^{2+2a}(mh_m)^a h_m^{1+a}[\mathrm{Var}^* U_{nm1}^*]^{1+a}}
$$

$$
\cdot \frac{2^{2+2a}\displaystyle\int_{-1}^{1} K^{2+2a}(u)\frac{\hat{f}_n(t-h_m u)}{\widehat{SG}_n^{1+2a}(t-h_m u)}\mathrm{d}u}{\varepsilon^{2+2a}(mh_m)^a\left[\displaystyle\int_{-1}^{1} K^2(u)\frac{\hat{f}_n(t-h_m u)}{\widehat{SG}_n(t-h_m u)}\mathrm{d}u - h_m\left(\int_{-1}^{1} K(u)\hat{f}_n(t-h_m u)\mathrm{d}u\right)^2\right]^{1+a}}
$$

$$
\to 0. \qquad (2.1.44)
$$

由(2.1.39),(2.1.40) 与(2.1.44) 知关于二重序列$\{U_{nmj}\}$ 的中心极限定理满足,于是 $mh_m \to \infty$ 时

$$
\frac{R_{1nm}^*}{\sqrt{\mathrm{Var}^* R_{1nm}^*}} \xrightarrow{\mathcal{L}^*} N(0,1). \qquad (2.1.45)
$$

又

$$
\mathrm{Var}^* R_{1nm}^* = mh_m^{-1}\left[\int_{-1}^{1} K^2(u)\frac{\hat{f}_n(t-h_m u)}{\widehat{SG}_n(t-h_m u)}\mathrm{d}u\right.
$$

$$
\left.-h_m\left(\int_{-1}^{1} K(u)\hat{f}_n(t-h_m u)\mathrm{d}u\right)^2\right], \qquad (2.1.46)
$$

再次利用引理 2.1.4(ii) 及(2.1.45),(2.1.46),即得 $mh_m \to \infty$ 时

$$
\sqrt{mh_m}R_{1nm}^* \xrightarrow{\mathcal{L}^*} N\left(0, \frac{\hat{f}_n(t)}{\widehat{SG}_n(t)}\int_{-1}^{1} K^2(u)\mathrm{d}u\right). \qquad (2.1.47)
$$

现取 $T_0$ 使 $t < T_0 < T_{\widehat{SH}_n}$,则 $m$ 充分大时 $t + h_m < T_0$,于是

$$
|R_{2nm}^*|
$$

$$
\leqslant \sup_{0\leqslant y\leqslant T_0} |\hat{G}_n^*(y) - \widehat{SG}_n(y)| \bar{G}_m^{*^{-1}}(T_0)\frac{1}{mh_m}\sum_{i=1}^{m}\frac{\delta_i^* K\left(\frac{t-z_i^*}{h_m}\right)}{\widehat{SG}_n(Z_i^*)},
$$

$$
(2.1.48)
$$

又

$$\frac{1}{mh_m}\sum_{i=1}^{m}\frac{\delta_i^* K\left(\frac{t-z_i^*}{h_m}\right)}{\overline{\widehat{SG}}_n(Z_i^*)}$$

$$=\frac{1}{h_m}\int_0^\infty \frac{K\left(\frac{t-y}{h_m}\right)}{\overline{\widehat{SG}}_n(y)}\mathrm{d}(\widehat{SH}_{nm1}(y)-\widetilde{SH}_{n1}(y))$$

$$+\frac{1}{h_m}\int_0^\infty \frac{K\left(\frac{t-y}{h_m}\right)}{\overline{\widehat{SG}}_n(y)}\mathrm{d}\,\widetilde{SH}_{n1}(y)\stackrel{\mathrm{d}}{=}I+II. \quad (2.1.49)$$

记 $A_{nm}=\{y:\widehat{SH}_{nm}(y)-\widehat{SH}_{nm}(y^-)>0\}$，$A_{nm}^c$ 为其补，则

$$|I|\leqslant\left|\int_{A_{nm}}\frac{h_m^{-1}K\left(\frac{t-y}{h_m}\right)}{\overline{\widehat{SG}}_n(y)}\mathrm{d}(\widehat{SH}_{nm}(y)-\widetilde{SH}_n(y))\right|$$

$$+\left|\int_{A_{nm}^c}\frac{h_m^{-1}K\left(\frac{t-y}{h_m}\right)}{\overline{\widehat{SG}}_n(y)}\mathrm{d}(\widehat{SH}_{nm}(y)-\widetilde{SH}_n(y)\right|$$

$$\leqslant 2\,\overline{\widehat{SG}}_n^{-1}(T_0)\sup_u K(u)h_m^{-1}\sup_{0\leqslant y\leqslant T_0}|\widehat{SH}_{nm}(y)-\widetilde{SH}_n(y)|.$$

由 $\widehat{SG}_n$ 的连续性及定理 2.1.4(a) 中的条件，即得

$$I\xrightarrow{a.s.^*}0,m\to\infty. \quad (2.1.50)$$

而由引理 2.1.4(i) 即得

$$II=\int_{-1}^{1}K(u)\hat{f}_n(t-h_mu))\mathrm{d}u\to\hat{f}_n(t),m\to\infty, \quad (2.1.51)$$

于是由(2.1.49)，(2.1.50) 与(2.1.51)，我们有

$$\frac{1}{mh_m}\sum_{i=1}^{m}\frac{\delta_i^* K\left(\frac{t-z_i^*}{h_m}\right)}{\widehat{SG}_n(Z_i^*)}\xrightarrow{a.s.^*}\hat{f}_n(t),m\to\infty.\quad(2.1.52)$$

在(2.1.48)中应用 Csörgö 与 Horváth[23] 中定理1与(2.1.52)即知 $h_m \log\log m\to 0$ 时

$$\sqrt{mh_m}R_{2nm}^*\xrightarrow{a.s.^*}0,m\to\infty.\quad(2.1.53)$$

而由定理2.1.4中条件 $mh_m^3\to 0$ 蕴涵 $h_m\log\log m\to 0$,即(2.1.53)成立.

由 $K$ 所满足的条件知 $\hat{f}_n'$ 存在且有界,于是

$$\begin{aligned}|R_{3nm}|&=\left|\frac{1}{h_m}\int_0^\infty K\left(\frac{t-y}{h_m}\right)\hat{f}_n(y)\mathrm{d}y-\hat{f}_n(t)\right|\\&=\left|\int_{-1}^{1}K(u)(\hat{f}_n(t-h_n u)-\hat{f}_n(t))\mathrm{d}u\right|\\&\leqslant\sup_t|\hat{f}_n'(t)|\left(\int_{-1}^{1}|u|K(u)\mathrm{d}u\right)h_m.\end{aligned}\quad(2.1.54)$$

由(2.1.53)即知 $mh_m^3\to 0$ 时

$$\sqrt{mh_m}R_{3nm}\to 0.\quad(2.1.55)$$

综合(2.1.36),(2.1.47),(2.1.53)与(2.1.55),由 Slutsky 定理即知,当 $mh_m\to\infty$,$mh_m^3\to 0$ 时

$$\sqrt{mh_m}(\hat{f}_{nm}^*-\hat{f}_n(t))\xrightarrow{\mathscr{L}^*}N\left(0,\frac{\hat{f}_n(t)}{\widehat{SG}_n(t)}\int_{-1}^{1}K^2(u)\mathrm{d}u\right),\quad(2.1.56)$$

即定理2.1.4(a)得证.

(2.1.56)与引理2.1.4一起立即证得(b),又完全仿(a)的证明可证(c). 最后由(b)与(c)即得(d)的证明. 至此,定理2.1.4得证.

注4:利用本文的方法证明 $\hat{f}_n(t)$ 的渐近正态性时,对 $f(t)$ 所加的条件比 Mielniczuk[56] 中要求 $f(t)$ 二阶有界导数这一条件弱.

### 2.1.4 Bootstrap 方差的几乎处处收敛性

**引理 2.1.5** 设对某 $\gamma > 0$,若取 $T_0$ 满足 $\widehat{SH}_n(T_0) > \gamma$,则当 $\frac{\sqrt{50}}{m\gamma} < \eta < \sqrt{2}$ 时

$$P(\sup_{0\leqslant y\leqslant T_0} |\bar{G}^*_{nm}(y) - \widehat{SG}_n(y)| > \eta) \leqslant \alpha\exp(-\lambda\gamma^2\eta^2 n), \tag{2.1.57}$$

其中 $\lambda$ 是正的绝对常数.

证明:证明与引理 2.1.2 的证明相同.

**定理 2.1.5** 在定理 2.1.4(b) 的条件下,对任取定的 $0 < t < \tau_H$,当 $h_m = m^{-\gamma}$(于是 $h_n = n^{-\gamma}$),$\frac{1}{3} < \gamma < 1, m \to \infty, n \to \infty$ 时,有

$$\hat{\sigma}^2_{nm} = \mathrm{Var}^*[\sqrt{mh_m}(\hat{f}^*_{nm}(t) - \hat{f}_n(t)] \xrightarrow{a.s.} \frac{f(t)}{G(t)}\int_{-1}^{1} K^2(u)\mathrm{d}u.$$

证明:我们先证当 $mh_m \to \infty, mh^3_m \to 0$ 时,有

$$\hat{\sigma}^2_{nm} = \mathrm{Var}^*[\sqrt{mh_m}(\hat{f}^*_{nm}(t) - \hat{f}_n(t)] \to \frac{\hat{f}_n(t)}{\widehat{SG}_n(t)}\int_{-1}^{1} K^2(u)\mathrm{d}u. \tag{2.1.59}$$

为证(2.1.59),由定理 2.1.4(a) 及 Chow 与 Teicher[22] 中 $P_{225}$ 推论 8 知,只须证 $mh_m(\hat{f}^*_{nm}(t) - \hat{f}_n(t))^2$ 一致可积,而它成立的一个充分条件是存在 $\delta > 0$,使

$$\sup_m E^* | \sqrt{mh_m}(\hat{f}^*_{nm} - \hat{f}_n(t)|^{2+\delta} < \infty, \tag{2.1.60}$$

注意到 $K(\cdot)$ 有界,因而 $|\hat{f}^*_{nm} - \hat{f}_n(t)| \leqslant \left(\frac{1}{h_m} + \frac{1}{h_n}\right)\sup_u K(u)$,由此即得

$$\begin{aligned} &E^* | \sqrt{mh_m}(\hat{f}^*_{nm} - \hat{f}_n(t)|^{2+\delta} \\ &= (2+\delta)\int_0^\infty x^{1+\delta}P^*(\sqrt{mh_m}\,|\hat{f}^*_{nm}(t) - \hat{f}_n(t)| > x)\mathrm{d}x \end{aligned}$$

$$\leqslant (2+\delta)\left[1+\int_1^{\sqrt{mh_m}\left(\frac{1}{h_m}+\frac{1}{h_n}\right)\sup_u K(u)} x^{1+\delta}\right.$$

$$\left.\times P^*(\sqrt{mh_m}\,|\hat{f}^*_{nm}(t)-\hat{f}_n(t)|>x)\right]\mathrm{d}x. \quad (2.1.61)$$

由(2.1.54)知当 $h_m = m^{-\gamma}, \frac{1}{3}<\gamma<1, m$ 充分大时对 $x\geqslant 1$

$$|R_{3nm}|<\frac{1}{3}\leqslant\frac{x}{3}, \quad (2.1.62)$$

因而由(2.1.36),知对 $x\geqslant 1$,有

$$P^*(\sqrt{mh_m}\,|\hat{f}^*_{nm}(t)-\hat{f}_n(t)|>x)$$

$$\leqslant P^*(\sqrt{mh_m}\,|R^*_{1nm}|>\frac{x}{3})+P^*(\sqrt{mh_m}\,|R^*_{2nm}|>\frac{x}{3})$$

$$\stackrel{\mathrm{d}}{=} D_{1nm}+D_{2nm}. \quad (2.1.63)$$

易得

$$D_{1nm}$$

$$=P^*\left(\left|\int_0^\infty \frac{K\left(\frac{t-y}{h_m}\right)}{\widehat{\widehat{SG}}_n(y)}\mathrm{d}(\widehat{SH}_{nm1}(y)-\widetilde{SH}_{n1}(y))\right|>\frac{1}{3}\sqrt{\frac{h_m}{m}}x\right)$$

$$\leqslant \alpha\frac{m^2}{h_m^2}x^{-4}E^*\left[\int_0^\infty \frac{K\left(\frac{t-y}{h_m}\right)}{\widehat{\widehat{SG}}_n(y)}\mathrm{d}(\widehat{SH}_{nm1}(y)-\widetilde{SH}_{n1}(y))\right]^4.$$

(2.1.64)

而

$$\Delta_{nm}\stackrel{\mathrm{d}}{=}E^*\left[\int_0^\infty \frac{K\left(\frac{t-y}{h_m}\right)}{\widehat{\widehat{SG}}_n(y)}\mathrm{d}(\widehat{SH}_{nm1}(y)-\widetilde{SH}_{n1}(y))\right]^4$$

$$=\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty\prod_{i=1}^4\frac{K\left(\frac{t-y_i}{h_m}\right)}{\widehat{\widehat{SG}}_n(y_i)}I[t-h_m\leqslant y_i\leqslant t+h_m]$$

$$\cdot \prod_{i=1}^{4} \mathrm{d}(\widehat{SH}_{nm1}(y) - \widetilde{SH}_{n1}(y)). \qquad (2.1.65)$$

记

$$h_n(y_1,y_2,y_3,y_4) = \prod_{i=1}^{4} \frac{K\left(\dfrac{t-y_i}{h_m}\right)}{\widehat{SG}_n(y_i)} I[t - h_m \leqslant y_i \leqslant t + h_m].$$

由于$\widehat{SG}_n$ 连续,因而 $m$ 充分大时$\overline{\widehat{SG}}_n(t+h_m) > \frac{1}{2}\overline{\widehat{SG}}_n(t)$,由此可推得

$$|h_n(y_1,y_2,y_3,y_4)| \leqslant 16 \sup_u K(u) \overline{\widehat{SG}}_n^{-4}(t), \qquad (2.1.66)$$

即对固定的 $t$ 与 $n$,$h_n(y_1,y_2,y_3,y_4)$ 有界,于是由(2.1.65),(2.1.66)与 Chang 与 Rao[16] 附录中引理得

$$\Delta_{nm} \leqslant \alpha \sup_u K(u) \widehat{SG} \stackrel{\mathrm{d}}{=}{}_n^{-4}(t) m^{-2}. \qquad (2.1.67)$$

而由(2.1.64),(2.1.65)与(2.1.66)即得

$$D_{1nm} \leqslant \alpha h_m^{-2} \overline{\widehat{SG}}_n^{-4}(t) x^{-4}. \qquad (2.1.68)$$

取 $T_0$,使 $t < T_0 < T_{\widehat{SH}_n}$,则当 $m$ 充分大$t + h_m < T_0 < T_{\widehat{SH}_n}$ 时,有

$$D_{2nm}$$

$$\leqslant P^*\left\{\frac{\sup\limits_{0\leqslant y\leqslant T_0} |\hat{G}_{nm}^*(y) - \widehat{SG}_n(y)|}{\bar{\hat{G}}_{nm}^*(T_0)} \frac{1}{\sqrt{mh_m}} \sum_{i=1}^{m} \frac{\delta_i^* K\left(\dfrac{t-z_i^*}{h_m}\right)}{\overline{\widehat{SG}}_n(Z_i^*)} > \frac{x}{3}\right\}$$

$$\leqslant P^*\left\{\frac{\sqrt{mh_m} \sup\limits_{0\leqslant y\leqslant T_0} |\hat{G}_{nm}^*(y) - \widehat{SG}_n(y)|}{\bar{\hat{G}}_{nm}^*(T_0)} \left| \frac{1}{mh_m} \sum_{i=1}^{m} \frac{\delta_i^* K\left(\dfrac{t-z_i^*}{h_m}\right)}{\overline{\widehat{SG}}_n(Z_i^*)} - \frac{1}{h_m}\int_0^\infty K\left(\frac{t-y}{h_m}\right) \mathrm{d}\,\widehat{SF}_n(y)\right| > \frac{x}{6}\right\}$$

$$
+ P^* \left( \frac{\sqrt{mh_m} \sup_{0 \leqslant y \leqslant T_0} |\hat{G}^*_{nm}(y) - \widehat{SG}_n(y)|}{\bar{\hat{G}}^*_{nm}(T_0)} \frac{1}{h_m} \int_0^\infty K\left(\frac{t-y}{h_m}\right) \mathrm{d}\, \widehat{SF}_n(y) > \frac{x}{6} \right)
$$

$$
\leqslant P^* \left( \frac{\sup_{0 \leqslant y \leqslant T_0} |\hat{G}^*_{nm}(y) - \widehat{SG}_n(y)|}{\bar{\hat{G}}^*_{nm}(T_0)} > \frac{1}{6} \right)
$$

$$
+ P^* \left( \sqrt{mh_m} \left| \frac{1}{mh_m} \sum_{i=1}^m \frac{\delta_i^* K\left(\frac{t-z_i^*}{h_m}\right)}{\overline{\widehat{SG}}_n(Z_i^*)} - \frac{1}{h_m} \int_0^\infty K\left(\frac{t-y}{h_m}\right) \mathrm{d}\, \widehat{SF}_n(y) \right| > x \right)
$$

$$
+ P^* \left( \frac{\sqrt{mh_m} \sup_{0 \leqslant y \leqslant T_0} |\hat{G}^*_{nm}(y) - \widehat{SG}_n(y)|}{\bar{\hat{G}}^*_{nm}(T_0)} > \frac{x}{6} \frac{1}{\frac{1}{h_m} \int_0^\infty K\left(\frac{t-y}{h_n}\right) \mathrm{d}\, \widehat{SF}_n(y)} \right)
$$

$$
\stackrel{\mathrm{d}}{=} E_{1nm} + E_{2nm} + E_{3nm}. \tag{2.1.69}
$$

易见

$$
E_{1nm} \leqslant P^* \left( \frac{\sup_{0 \leqslant y \leqslant T_0} |\hat{G}^*_{nm}(y) - \widehat{SG}_n(y)|}{\overline{\widehat{SG}}_n(T_0)} > \frac{1}{12} \right)
$$

$$
+ P^* \left( \bar{\hat{G}}^*_{nm}(T_0) < \frac{1}{2} \overline{\widehat{SG}}_n(T_0) \right). \tag{2.1.70}
$$

由(2.1.70),并使用下面事实

$$P^*\left(\bar{G}^*_{nm}(T_0) < \frac{1}{2}\overline{\widehat{SG}}_n(T_0)\right)$$

$$\leqslant P^*\left(|\bar{G}^*_{nm}(T_0) - \overline{\widehat{SG}}_n(T_0)| > \frac{1}{2}\overline{\widehat{SG}}_n(T_0)\right)$$

及引理 2.1.5(在引理 2.1.5 中取 $\gamma = \frac{1}{2}\overline{\widehat{SH}}_n(T_0)$),即得

$$E_{1nm} \leqslant \alpha\cdot\exp\{-Cm\overline{\widehat{SG}}{}^2_n(T_0)\overline{\widehat{SH}}{}^2_n(T_0)\}. \quad (2.1.71)$$

由 $E_{2nm}$ 与 $R^*_{1nm}$ 的定义,利用(2.1.68),得

$$E_{2nm} = P^*(\sqrt{mh_m}|R^*_{1nm}| > x) \leqslant Ch_m^{-2}\overline{\widehat{SG}}{}_n^{-4}x^{-4}. \quad (2.1.72)$$

由引理 2.1.3(i) 知 $m$ 充分大时 $\left|\frac{1}{h_m}\int_0^\infty K\left(\frac{t-y}{h_m}\right)\mathrm{d}\widehat{SF}_n(y)\right| \leqslant 2\hat{f}_n(t)$,且因 $x\geqslant 1$,因此

$$E_{3nm} \leqslant P^*\left(\frac{\sup\limits_{0\leqslant y\leqslant T_0}|\hat{G}^*_{nm}(y) - \widehat{SG}_n(y)|}{\bar{G}^*_{nm}(T_0)} > \frac{1}{12\hat{f}_n(t)\sqrt{mh_m}}\right).$$

于是类似于(2.1.71)的证明可得

$$E_{3nm} \leqslant \alpha\cdot\exp\{-\alpha\overline{\widehat{SG}}{}^2_n(T_0)\overline{\widehat{SH}}{}^2_n(T_0)\min\{\hat{f}_n^{-2},1\}h_m^{-1}\}. \quad (2.1.73)$$

综合(2.1.69),(2.1.71)—(2.1.73)即得

$$D_{2nm} \leqslant \alpha[\exp\{-\alpha\overline{\widehat{SG}}{}^2_n(T_0)\overline{\widehat{SH}}{}^2_n(T_0)\min\{\hat{f}_n^{-2},1\}h_m^{-1}\} + h_m^{-2}\overline{\widehat{SG}}{}_n^{-4}(t)x^{-4}]. \quad (2.1.74)$$

再综合(2.1.63),(2.1.68)与(2.1.74),得

$$P^*(\sqrt{mh_m}|\hat{f}^*_{nm} - \hat{f}_n(t)| > x)$$
$$\leqslant \alpha[\exp\{-\alpha\overline{\widehat{SG}}{}^2_n(T_0)\overline{\widehat{SH}}{}^2_n(T_0)\min\{\hat{f}_n^{-2},1\}h_m^{-1}\} + h_m^{-2}\overline{\widehat{SG}}{}_n^{-4}(t)x^{-4}], \quad x\geqslant 1. \quad (2.1.75)$$

于是(2.1.61)与(2.1.75)即得

$$
\begin{aligned}
&E^* \sqrt{mh_m}\,|\hat{f}^*_{nm} - \hat{f}_n(t)|^{2+\delta} \\
&\leqslant (2+\delta)\left[1 + \frac{\alpha}{2+\delta}\left(\frac{m}{h_m}\right)^{1+\frac{\delta}{2}}\right. \\
&\exp\{-\alpha\,\widehat{SG}_n^2(T_0)\,\overline{SH}_n^2(T_0)\min\{\hat{f}_n^{-2},1\}h_m^{-1}\} \\
&\left. + m^{-1-\frac{\delta}{2}} h_m^{-3+\frac{\delta}{2}}\,\widehat{SG}_n^{-4}(t)\right].
\end{aligned} \tag{2.1.76}
$$

于是取 $h_m = m^{-\gamma}, \frac{1}{3} < \gamma < 1, 0 < \delta < \frac{2(1-\gamma)}{1+\gamma}$ 时

$$
\sup_m E^* \mid \sqrt{mh_m}(\hat{f}^*_{nm}(t) - \hat{f}_n(t))|^{2+\delta} < \infty.
$$

这就证明了(2.1.60),因而(2.1.59)成立,最后利用引理2.1.4即得定理2.1.5的证明.

## §2.2 失效率估计的一些大样本性质

### 2.2.1 记号与定义

设 $T_1, T_2, \cdots, T_n$ 是独立同分布非负的随机变量,有共同的分布函数 $F$ 和相应的连续概率密度 $f$, $C_1, C_2, \cdots, C_n$ 是独立同分布表示删失的随机变量,有共同的连续分布函数 $G$,且诸 $T_i$ 独立于诸 $C_i$. 在随机右删失模型中,我们观察到的数据为 $(X_i, \delta_i)$,其中 $X_i = \min(T_i, C_i)$, $\delta_i = I[T_i \leqslant C_i]$, $i = 1, 2, \cdots, n$,且 $I[\cdot]$ 表示某事件的示性函数,易见 $X_1, X_2, \cdots, X_n$ iid 具有连续分布函数 $H = 1 - (1-F)(1-G)$.

本节所要考虑的问题是利用上面删失数据构造失效率 $\lambda(t) = \frac{f(t)}{1-F(t)}$ 的估计并研究它的一些大样本性质.

如引言中所述,关于失效率估计,由于其实用的重要性,引起了许多统计学者的注意,他们根据 $\lambda(t)$ 的各种不同的表示构造了 $\lambda(t)$ 的不同估计. 本节根据 $\lambda(t)$ 的如下表示:

$$\lambda(t) = \frac{\mathrm{d}}{\mathrm{d}t}[-\log(1 - F(t))]$$

定义 $\lambda(t)$ 的估计如下：

$$\lambda_n(t) = h_n^{-1}\int K\left(\frac{t-s}{h_n}\right)\frac{\mathrm{d}\hat{F}_n(s)}{1-\hat{F}_n(s)},$$

其中$\{h_n\}$是以某一速度趋于0的常数序列，$K$是具有有界支撑的概率密度核函数，$\hat{F}_n$是按下面方式定义的KM估计

$$1 - \hat{F}_n(t) = \begin{cases} \prod_{j=1}^{n}\left(\frac{N(X_j)}{N(X_j)+1}\right)^{I[X_j \leqslant t, \delta_j = 1]}, & t \leqslant \max\limits_{1\leqslant j\leqslant n} X_j, \\ 0, & \text{其他}, \end{cases}$$

$$N(\cdot) = \#\{X_j > \cdot\} = \sum_{i=1}^{n} I[X_i > \cdot].$$

本节约定$0^0 = 1$，$C$表示任何所需的可能与固定$t$有关的常数. 设

$$H_1(t) = P(X_1 \leqslant t,\ \delta_1 = 1),$$

$$H_{n1}(t) = n^{-1}\sum_{i=1}^{n} I[X_i \leqslant t,\ \delta_i = 1],$$

$$H_0(t) = P(X_1 \leqslant t, \delta_i = 0),$$

$$H_{n0} = \frac{1}{n}\sum_{i=1}^{n} I[X_i \leqslant t, \delta_i = 0],$$

$$H_n(t) = n^{-1}\sum_{i=1}^{n} I[X_i \leqslant t],$$

且对任何分布函数 $W$，记 $\overline{W}(\cdot) = 1 - W(\cdot)$. 下面，我们固定某点 $0 < \tau < \infty$ 使得 $\overline{H}(\tau) > 0$.

为陈述我们的定理简洁，特列下面供以后选择使用的条件.

$C1K$：$K(\cdot)$是具有有界支撑集$(p,q)$的有界变差概率密度核函数，其中 $-\infty < p < q < +\infty$.

$C2\lambda$：$\lambda(t)$在任何区间$[0,\tau']$上存在$k$阶有界的导数.

$C3\lambda$：$\lambda(t)$在$[0,\infty)$上连续.

$C4H$:$H$ 有概率密度 $h(t)$，且 $h(t)$ 在任何区间 $[0,\tau']$ 上有界.

本节研究 $\lambda_n(t)$ 的弱收敛与强一致性收敛速度，给出 $\lambda_n(t)$ 的渐近表示，作为该表示的一个应用，我们使用它证明 $\lambda_n(t)$ 的渐近正态性.

### 2.2.2 $\lambda_n(t)$ 的弱收敛速度

**定理 2.2.1** 若 $C1K$, $C2\lambda$ 满足，则存在均值为 0 的 Gaussian 过程 $G_n(t)$，使得对任何 $0 < t_1 < t_2 < \tau$，有

$$\sup_{t_1 \leqslant t \leqslant t_2} | \sqrt{n} h_n (\lambda_n(t) - \lambda(t)) - G_n(t) |$$

$$= O(n^{-\frac{1}{2}} \log n) + O(\sqrt{n} h_n^{k+1}), \quad a.s.$$

且

$$EG_n(s)G_n(t)$$

$$= h_n \lambda(s) \int_0^\infty K\left(\frac{t-s}{h_n} + v\right) K(v) \mathrm{d}v, \quad 0 \leqslant s \leqslant t < \infty.$$

为了证明定理 2.2.1，我们必须证明下面引理 2.2.1 与引理 2.2.2.

**引理 2.2.1** 以概率 1，我们有

$$\frac{1}{1 - H_n(t)}$$

$$= \frac{1}{1 - H(t)} \left[ 1 + \frac{H_n(t) - H(t)}{1 - H(t)} + R\left(\frac{H(t) - H_n(t)}{1 - H(t)}\right) \right],$$

而

$$R\left(\frac{H(t) - H_n(t)}{1 - H(t)}\right) \leqslant 2\left(\frac{H_n(t) - H(t)}{1 - H(t)}\right)^2$$

对 $t \in [0, \tau]$ 成立.

证明：由 Glivenko-Cantelli 引理，一致地有

$$\frac{1}{2}\overline{H}(t) \leqslant \overline{H}_n(t) \leqslant \frac{3}{2}\overline{H}(t), \quad a.s. \qquad (2.2.1)$$

在 $t \in [0,\tau]$ 上成立，即对 $t \in [0,\tau]$，我们一致地有

$$\left|\frac{H_n(t)-H(t)}{1-H(t)}\right| \leqslant \frac{1}{2}, \quad a.s. \tag{2.2.2}$$

注意到$\frac{1}{1+x} = 1 - x + R(x)$，其中 $R(x) \leqslant 2x^2$，若 $|x| \leqslant \frac{1}{2}$. 因此由(2.2.2)，我们一致地有

$$\begin{aligned}
&\frac{1}{1-H_n(t)} \\
&= \frac{1}{(1-H(t))\left[1+\frac{H(t)-H_n(t)}{1-H(t)}\right]} \\
&= \frac{1}{1-H(t)}\left[1+\frac{H_n(t)-H(t)}{1-H(t)} + R\left(\frac{H(t)-H_n(t)}{1-H(t)}\right)\right].
\end{aligned}$$

而

$$R\left(\frac{H(t)-H_n(t)}{1-H(t)}\right) \leqslant 2\left(\frac{H_n(t)-H(t)}{1-H(t)}\right)^2, \quad a.s.$$

在 $t \in [0,\tau]$ 时成立. 这就完成了引理 2.2.1 的证明.

**引理 2.2.2** 存在均值为 0 的 Gaussian 过程 $G_n(t)$ 使得

$$\begin{aligned}
&\sup_{0\leqslant t\leqslant q\tau} \left|\sqrt{n}(\log\hat{\bar{F}}_n(t) - \log\bar{F}(t)) - G_n(t)\right| \\
&= O(n^{-\frac{1}{2}}\log n). \quad a.s.
\end{aligned} \tag{2.2.3}$$

证明：对 $t \in [0,\tau]$，

$$\begin{aligned}
\log\hat{\bar{F}}_n &= \sum_{i=1}^n I[X_i \leqslant t, \delta_i = 1]\log\left[1 - \frac{1}{N(X_i)+1}\right] \\
&= -\sum_{i=1}^n I[X_i \leqslant t, \delta_i = 1]\sum_{j=1}^\infty \frac{1}{j(N(X_i)+1)^j} \\
&= -\sum_{i=1}^n I[X_i \leqslant t, \delta_i = 1]\sum_{j=1}^\infty \frac{1}{j(n - nH_n(X_i)+1)^j} \\
&= -n\int_0^t \frac{\mathrm{d}\widetilde{F}_n(s)}{n - nH_n(s) + 1}
\end{aligned}$$

$$- n\int_0^t\left(\sum_{j=2}^{\infty} \frac{1}{j(n - nH_n(s) + 1)^j}\right)\mathrm{d}\widetilde{F}_n(s). \quad (2.2.4)$$

注意到

$$\log\overline{F} = -\int_0^t \frac{\mathrm{d}\widetilde{F}(s)}{1 - H(s)}. \quad (2.2.5)$$

由(2.2.4),(2.2.5),我们有

$$\log\hat{\overline{F}}_n(t) - \log\overline{F}(t) = I_{11n} + I_{12n}, \quad (2.2.6)$$

其中

$$\begin{cases} I_{11n}(t) = -\left[n\int_0^t \dfrac{\mathrm{d}\widetilde{F}_n(s)}{n - nH_n(s) + 1} - \int_0^t \dfrac{\mathrm{d}\widetilde{F}(s)}{1 - H(s)}\right], \\ I_{12n}(t) = -n\int_0^t\left(\sum_{j=2}^{\infty} \dfrac{1}{j(n - nH_n(s) + 1)^j}\right)\mathrm{d}\widetilde{F}_n(s). \end{cases} \quad (2.2.7)$$

易见

$$\sum_{j=2}^{\infty} \frac{1}{j(n - nH_n(s) + 1)^j} \leqslant \frac{3}{2}\frac{1}{(n - nH_n(s) + 1)^2}. \quad (2.2.8)$$

由(2.2.8) 与引理 2.2.1,我们有

$$\sup_{0\leqslant t\leqslant \tau} |I_{12n}(t)| = O(n^{-1}). \quad (2.2.9)$$

易见

$$\begin{aligned} I_{11n} = -\Bigg[ & n\left(\int_0^t \frac{\mathrm{d}(H_{n1}(s) - H_1(s))}{n - nH_n(s) + 1} - \int_0^t \frac{\mathrm{d}H_{n1}(s) - H_1(s)}{n - nH_n(s)}\right) \\ & + \int_0^t \frac{\mathrm{d}(H_{n1}(s) - H_1(s))}{1 - H_n(s)} \\ & + n\left(\int_0^t \frac{\mathrm{d}H_1(s)}{n - nH_n(s) + 1} - \int_0^t \frac{\mathrm{d}H_1(s)}{n - nH_n(s)}\right) \\ & + \left(\int_0^t \frac{\mathrm{d}H_1(s)}{1 - H_n(s)} - \int_0^t \frac{\mathrm{d}H_1(s)}{1 - H(s)}\right)\Bigg] \end{aligned}$$

$$
\begin{aligned}
&=\int_0^t \frac{\mathrm{d}(H_{n1}(s)-H_1(s))}{(n-nH_n(s)+1)(1-H_n(s))} \\
&\quad -\int_0^t \frac{\mathrm{d}(H_{n1}(s)-H_1(s))}{1-H_n(s)} \\
&\quad +\int_0^t \frac{\mathrm{d}H_1(s)}{(n-nH_n(s)+1)(1-H_n(s))} \\
&\quad -\int_0^t \frac{H_n(s)-H(s)}{(1-H_n(s))(1-H(s))}\mathrm{d}H_1(s) \\
&\stackrel{\mathrm{d}}{=} I_{21n}-I_{22n}+I_{23n}-I_{24n}.
\end{aligned} \tag{2.2.10}
$$

使用引理 2.2.1,得到

$$
\begin{cases}
\sup\limits_{0\leqslant t\leqslant \tau}|I_{21n}(t)| = O(n^{-1}), \\
\sup\limits_{0\leqslant t\leqslant \tau}|I_{23n}(t)| = O(n^{-1}).
\end{cases} \tag{2.2.11}
$$

由 Major 与 Rejtö[54] 知,存在 Brownian 桥 $B(y),0\leqslant y\leqslant 1$ 使得

$$
\begin{aligned}
&\sup_{0\leqslant s<+\infty}|\sqrt{n}(H_{n1}(s)-H_1(s))-B(H_1(s))| \\
&= O(n^{-\frac{1}{2}}\log n), \quad a.s.
\end{aligned} \tag{2.2.12}
$$

$$
\begin{aligned}
&\sup_{0\leqslant s<\infty}|\sqrt{n}(H_{n0}(s)-H_0(s))-B(1-H_0(s))| \\
&= O(n^{-\frac{1}{2}}\log n), \quad a.s.
\end{aligned} \tag{2.2.13}
$$

定义 Gaussian 过程

$$
G_1(t)=\int_0^t \frac{B(H_1(s))+B(1-H_0(s))}{(1-H(s))^2}\mathrm{d}H_1(s).
$$

由(2.2.12),(2.2.13),引理 2.2.1,Glivenko-Cantelli 引理与下面事实:

$$
H_n(s)-H(s)=H_{n1}(s)-H_1(s)+H_{n0}(s)-H_0(s), \tag{2.2.14}
$$

我们有

$$\sup_{0\leqslant t\leqslant \tau}|\sqrt{n}I_{24n}(t)-G_1(t)|$$

$$\leqslant \sup_{0\leqslant t\leqslant \tau}\int_0^t \frac{|\sqrt{n}(H_n(s)-H(s))-B(H_1(s))-B(1-H_0(s))|}{(1-H(s))^2}\mathrm{d}H_1(s)$$

$$+\sqrt{n}\sup_{0\leqslant t\leqslant \tau}\int_0^t \frac{(H_n(s)-H(s))^2}{(1-H(s))^3}\mathrm{d}H_1(s)$$

$$+2\sqrt{n}\sup_{0\leqslant t\leqslant \tau}\int_0^t \frac{(H_n(s)-H(s))^3}{(1-H(s))^4}\mathrm{d}H_1(s)$$

$$=O(n^{-\frac{1}{2}}\log n),\quad a.s. \tag{2.2.15}$$

再定义 Gaussian 过程

$$G_2(t)=\int_0^t \frac{\mathrm{d}B(H_1(s))}{1-H(s)},$$

使用引理 2.2.1, Glivenko-Cantelli 引理(2.2.12), 类似于(2.2.15)我们有

$$\sup_{0\leqslant t\leqslant \tau}|\sqrt{n}I_{22n}(t)-G_2(t)|=O(n^{-\frac{1}{2}}\log n),\quad a.s. \tag{2.2.16}$$

又定义 Gaussian 过程

$$G^*(t)=-G_1(t)-G_2(t). \tag{2.2.17}$$

由(2.2.6),(2.2.10),(2.2.11),(2.2.15),(2.2.16)与(2.2.17),我们得到

$$\sup_{0\leqslant t\leqslant \tau}|\sqrt{n}(\log\hat{\bar{F}}_n(t)-\log\bar{F}(t))-G^*(t)|$$

$$=O(n^{-\frac{1}{2}}\log n),\quad a.s. \tag{2.2.18}$$

这就证明了引理 2.2.2.

定理 2.2.1 的证明:我们看到

$$\lambda_n(t)-\lambda(t)$$

$$=\left(-h_n^{-1}\int_{t-ph_n}^{t-qh_n}K\left(\frac{t-s}{h_n}\right)\mathrm{d}\log\hat{\bar{F}}_n\right.$$

$$+ h_n^{-1}\int_{t-ph_n}^{t-qh_n} K\left(\frac{t-s}{h_n}\right)\mathrm{d}\log\bar{F}(s)\Big)$$

$$+ \left(h_n^{-1}\int_{t-ph_n}^{t-qh_n} K\left(\frac{t-s}{h_n}\right)\lambda(s)\mathrm{d}s - \lambda(t)\right)$$

$$\stackrel{\mathrm{d}}{=} I_{31n} + I_{32n} \tag{2.2.19}$$

对任何 $0 < t < \tau$ 成立. 分部积分后,我们有

$$I_{31n} = - h_n^{-1}\int_{t-ph_n}^{t-qh_n} K\left(\frac{t-s}{h_n}\right) \mathrm{d}[\log\hat{\bar{F}}_n(s) - \log\bar{F}(s)]$$

$$= h_n^{-1}\int_{t-ph_n}^{t-qh_n} (\log\hat{\bar{F}}_n(s) - \log\bar{F}(s))\mathrm{d}K\left(\frac{t-s}{h_n}\right).$$

现定义 Gaussian 过程

$$G_n(t) = \int_{t-ph_n}^{t-qh_n} G^*(s)\mathrm{d}K\left(\frac{t-s}{h_n}\right). \tag{2.2.20}$$

由引理 2.2.2,得到

$$\sup_{t_1\leqslant t\leqslant t_2} |\sqrt{n}h_n I_{31n} - G_n(t)| = O(n^{-\frac{1}{2}}\log n), \quad a.s. \tag{2.2.21}$$

对任何 $0 < t_1 < t_2 < \tau$ 成立.

使用 Wang[84] 中(2.9) 的证明方法,可证

$$\sup_{t_1\leqslant t\leqslant t_2} |I_{32n}| \leqslant C\cdot h_n^k \tag{2.2.22}$$

对任何 $0 < t_1 < t_2 < \tau$ 成立.

(2.2.19),(2.2.21) 与(2.2.22) 一起就证明了

$$\sup_{t_1\leqslant t\leqslant t_2} |\sqrt{n}h_n(\lambda_n(t) - \lambda(t)) - G_n(t)|$$

$$= O(n^{-\frac{1}{2}}\log n) + O(\sqrt{n}h_n^{k+1}), \quad a.s.$$

因此,定理 2.2.1 得证.

**推论 2.2.1** 在定理 2.2.1 的条件下,若 $h_n = n^{-\alpha}$ 且 $\frac{1}{k+1} <$

$\alpha < \frac{1}{2}$,我们有

$$\sup_{t_1 \leqslant t \leqslant t_2} | \sqrt{n} h_n (\lambda_n(t) - \lambda(t)) - G_n(t) |$$
$$= O(n^{-\frac{1}{2}} \log n), \quad a.s.$$

对任何 $0 < t_1 < t_2 < \tau$ 成立,其中 $G_n(t)$ 如(2.2.20)所定义.

### 2.2.3 $\lambda_n(t)$ 的强一致相合性

**引理 2.2.3** 以概率 1,我们有

$$\sup_{0 \leqslant t \leqslant \tau} | \log \bar{\hat{F}}_n(t) - \log \bar{F}(t) | = O(n^{-\frac{1}{2}} \sqrt{\log\log n}).$$

证明:由(2.2.6),(2.2.7),(2.2.9),(2.2.10)与(2.2.11),我们有

$$\log \bar{\hat{F}}_n(t) - \log \bar{F}(t)$$
$$= -\int_0^t \frac{d(H_{n1}(s) - H_1(s))}{1 - H_n(s)}$$
$$- \int_0^t \frac{H_n(s) - H(s)}{(1 - H_n(s))(1 - H(s))} dH_1(s) + O(n^{-1}) \quad (2.2.23)$$

对任何 $0 \leqslant t \leqslant \tau$ 成立.

在(2.2.23)中使用引理 2.2.1 与 Glivenko-Cantelli 引理,可得

$$\sup_{0 \leqslant t \leqslant \tau} | \log \bar{\hat{F}}_n(t) - \log \bar{F}(t) | = O(n^{-\frac{1}{2}} \sqrt{\log\log n}). \quad a.s.$$

于是,引理 2.2.3 得证.

**定理 2.2.2** 在条件 $C1K, C3\lambda$,下,若 $h_n^{-1} n^{-\frac{1}{2}} \sqrt{\log\log n} \to 0$,我们有

$$\sup_{0 \leqslant t \leqslant t_0} | \lambda_n(t) - \lambda(t) | \xrightarrow{a.s.} 0$$

对任何 $0 < t_0 < \tau$ 成立.

证明:易见

$$
\begin{aligned}
|\lambda_n(t)-\lambda(t)| \leqslant & h_n^{-1}\left|\int_{t-ph_n}^{t-qh_n} K\left(\frac{t-s}{h_n}\right) \mathrm{dlog}\hat{\bar{F}}_n(s)\right. \\
& \left. -\int_{t-ph_n}^{t-qh_n} K\left(\frac{t-s}{h_n}\right) \mathrm{dlog}\bar{F}(s)\right| \\
& + \left|h_n^{-1}\int_{t-ph_n}^{t-qh_n} K\left(\frac{t-s}{h_n}\right)\lambda(s)\mathrm{d}s-\lambda(t)\right)\Bigg| \\
\stackrel{\mathrm{d}}{=} & \amalg_{11n}+\amalg_{12n} \qquad (2.2.24)
\end{aligned}
$$

对任何 $0 \leqslant t \leqslant t_0 < \tau$ 成立. 分部积分并使用条件 $C1K$,可得

$$
\amalg_{11n} \leqslant h_n^{-1} \sup_{0\leqslant s\leqslant \tau} |\log\hat{\bar{F}}_n(s)-\log\bar{F}(s)| \int_p^q |\mathrm{d}K(u)|
$$

对充分大的 $n$ 与 $0 < t < t_0$ 成立. 由引理2.2.3且注意到 $K(\cdot)$ 是 $(p,q)$ 上的有界变差核函数,于是我们有

$$
\sup_{0\leqslant t\leqslant t_0} |\amalg_{11n}| = O(h_n^{-1}n^{-\frac{1}{2}}\sqrt{\log\log n}), \quad a.s. \tag{2.2.25}
$$

观察到

$$
\begin{aligned}
\amalg_{12n} &= \left|h_n^{-1}\int_{t-ph_n}^{t-qh_n} K\left(\frac{t-s}{h_n}\right)\lambda(s)\mathrm{d}s-\int_p^q K(u)\lambda(t)\mathrm{d}u\right| \\
&= \left|\int_p^q K(u)\lambda(t-h_nu)\mathrm{d}u-\int_p^q K(u)\lambda(t)\mathrm{d}u\right|. \qquad (2.2.26)
\end{aligned}
$$

已知 $\lambda(t)$ 是连续函数,因此对任何 $\varepsilon > 0$,关于 $0 \leqslant t \leqslant \tau$, $p \leqslant u \leqslant q$ 我们一致地有

$$
|\lambda(t-h_nu)-\lambda(t)| < \varepsilon. \tag{2.2.27}
$$

由(2.2.26),(2.2.27),知

$$
\sup_{0\leqslant t\leqslant \tau} |\amalg_{12n}| < \varepsilon \tag{2.2.28}
$$

对充分大的 $n$ 成立.

因此,综合(2.2.24),(2.2.25)与(2.2.28)我们就得定理2.2.2的证明.

**定理 2.2.3** 在条件 $C1K$, $C2\lambda$ 下,以概率1,对任何 $0 < t_1$

$< t_2 < \tau$,有

$$\sup_{t_1 \leqslant t \leqslant t_2} |\lambda_n(t) - \lambda(t)| = O(h_n^{-1} n^{-\frac{1}{2}} \sqrt{\log\log n}) + O(h_n^k).$$

证明:应用(2.2.19),且注意到$II_{12n} = |I_{32n}|$,此处$I_{32n}$如(2.2.19)所定义.由(2.2.22)我们有

$$\sup_{t_1 \leqslant t \leqslant t_2} II_{12n} = O(h_n^k) \qquad (2.2.29)$$

对任何$0 < t_1 < t_2 < \tau$成立.因此,(2.2.24),(2.2.25)与(2.2.29)一起就证明了定理2.2.3.

### 2.2.4 $\lambda_n(t)$的渐近表示与渐近正态性

记

$$\tilde{\lambda}_n(t) = h_n^{-1} \int K\left(\frac{t-s}{h_n}\right) \frac{\mathrm{d}F(s)}{1-F(s)}. \qquad (2.2.30)$$

**定理 2.2.4** 在条件 $C1K$ 与 $C4H$ 下,以概率1有

$$\lambda_n(t) - \tilde{\lambda}_n(t) = \frac{\tilde{f}_n(t) - E\tilde{f}_n(t)}{1-H(t)} + O((nh_n)^{-1}\log\log n) + O(n^{-\frac{1}{2}}\sqrt{\log\log n})$$

对任何$0 < t < \tau$成立,其中

$$\tilde{f}_n(t) = h_n^{-1} \int K\left(\frac{t-s}{h_n}\right) \mathrm{d}H_{n1}(s). \qquad (2.2.31)$$

为证定理2.2.4,我们必须证明下面的引理.

**引理 2.2.4** 对任何$0 \leqslant t \leqslant \tau$,我们有

$$\begin{aligned}
&\log\bar{\hat{F}}_n - \log\bar{F} \\
&= -\frac{H_{n1}(t) - H_1(t)}{1-H(t)} + \int_0^t \frac{H_{n1}(s) - H_1(s)}{(1-H(s))^2} \mathrm{d}H(s) \\
&\quad + \int_0^t \frac{H_n(s) - H(s)}{(1-H_n(s))(1-H(s))} \mathrm{d}H_{n1}(s) \\
&\quad - n^{-1} \int_0^t \frac{\mathrm{d}H_{n1}(s)}{(1-H_n(s)+\frac{1}{n})(1-H_n(s))} + I_{12n}(t), \qquad (2.2.32)
\end{aligned}$$

其中 $I_{12n}(t)$ 如(2.2.7)所定义.

证明:由(2.2.11),(2.2.12)并分部积分可知对任何 $0 \leqslant t < \tau$,我们有

$$
\begin{aligned}
&\log \bar{\hat{F}}_n - \log \bar{F} \\
=& -\left[n\int_0^t \frac{\mathrm{d}H_{n1}(s)}{n - nH_n(s) + 1} - \int_0^t \frac{\mathrm{d}H_{n1}(s)}{1 - H_n(s)}\right. \\
&+ \int_0^t \frac{\mathrm{d}H_{n1}(s)}{1 - H_n(s)} - \int_0^t \frac{\mathrm{d}H_{n1}(s)}{1 - H(s)} \\
&\left.+ \int_0^t \frac{\mathrm{d}H_{n1}(s)}{1 - H(s)} - \int_0^t \frac{\mathrm{d}H_1(s)}{1 - H(s)}\right] + I_{12n}(t) \\
=& -\int_0^t \frac{\mathrm{d}(H_{n1}(s) - H_1(s))}{1 - H(s)} \\
&+ n^{-1}\int_0^t \frac{\mathrm{d}H_{n1}(s)}{(1 - H_n(s) + \frac{1}{n})(1 - H_n(s))} \\
&- \int_0^t \frac{H_n(s) - H(s)}{(1 - H_n(s))(1 - H(s))}\mathrm{d}H_{n1}(s) + I_{12n}(t) \\
=& -\frac{H_{n1}(t) - H_1(t)}{1 - H(t)} + \int_0^t \frac{H_{n1}(s) - H_1(s)}{(1 - H(s))^2}\mathrm{d}H(s) \\
&+ n^{-1}\int_0^t \frac{\mathrm{d}H_{n1}(s)}{(1 - H_n(s) + \frac{1}{n})(1 - H_n(s))} \\
&- \int_0^t \frac{H_n(s) - H(s)}{(1 - H_n(s))(1 - H(s))}\mathrm{d}H_{n1}(s) + I_{12n}(t). \quad (2.2.33)
\end{aligned}
$$

这就证明了引理2.2.4

定理2.2.4的证明:分部积分并使用引理2.2.3,对任何固定的 $0 < t < \tau$,我们得

$$
\begin{aligned}
&\lambda_n(t) - \tilde{\lambda}(t) \\
=& -h_n^{-1}\int_{t-ph_n}^{t-qh_n} K\left(\frac{t-s}{h_n}\right)\mathrm{d}(\log \bar{\hat{F}}_n(s) - \log \bar{F}(s))
\end{aligned}
$$

$$= h_n^{-1}\int_{t-ph_n}^{t-qh_n}(\log\bar{\hat{F}}_n(s) - \log\bar{F}(s))\mathrm{d}K\left(\frac{t-s}{h_n}\right)$$

$$= - h_n^{-1}\int_{t-ph_n}^{t-qh_n}\frac{H_{n1}(s) - H_1(s)}{1 - H(s)}\mathrm{d}K\left(\frac{t-s}{h_n}\right)$$

$$+ h_n^{-1}\int_{t-ph_n}^{t-qh_n}\int_0^s\frac{H_{n1}(y) - H_1(y)}{(1 - H(y))^2}\mathrm{d}H(y)\mathrm{d}K\left(\frac{t-s}{h_n}\right)$$

$$+ (nh_n)^{-1}\int_{t-ph_n}^{t-qh_n}\int_0^s\frac{\mathrm{d}H_{n1}(y)}{(1 - H_n(y) + \frac{1}{n})(1 - H_n(y))}\mathrm{d}K\left(\frac{t-s}{h_n}\right)$$

$$- h_n^{-1}\int_{t-ph_n}^{t-qh_n}\int_0^s\frac{H_n(y) - H(y)}{(1 - H_n(y))(1 - H(y))}\mathrm{d}H_{n1}(y)\mathrm{d}K\left(\frac{t-s}{h_n}\right)$$

$$+ h_n^{-1}\int_{t-ph_n}^{t-qh_n}I_{12n}(s)\mathrm{d}K\left(\frac{t-s}{h_n}\right)$$

$$\stackrel{\mathrm{d}}{=} - \Delta_{1n} + \Delta_{2n} + \Delta_{3n} - \Delta_{4n} + \Delta_{5n}. \tag{2.2.34}$$

令 $u = \dfrac{t-s}{h_n}$,且通过分部积分,我们有

$$\Delta_{2n} = \int_p^q K(u)\frac{\widetilde{F}_n(t - h_n u) - \widetilde{F}(t - h_n)}{(1 - H(t - h_n u))^2}h(t - h_n u)\mathrm{d}u.$$

由此及(2.2.14)与 Kiefer 定理(见 Kiefer[46]),得

$$|\Delta_{2n}| \leqslant \sup_{0\leqslant t\leqslant\tau}|\widetilde{F}_n(t) - \widetilde{F}(t)|(1 - H(\tau))^{-2}\sup_{0\leqslant t\leqslant\tau}h(t)\int_p^q K(u)\mathrm{d}u$$

$$\leqslant Cn^{-\frac{1}{2}}\sqrt{\log\log n}, \quad a.s. \tag{2.2.35}$$

而使用 2.2.1,我们得到

$$|\Delta_{3n}| \leqslant (nh_n)^{-1}(1 - H(\tau))^{-2}\int_p^q|\mathrm{d}K(u)| \leqslant C(nh_n)^{-1}. \tag{2.2.36}$$

观察到

$$\Delta_{4n} = h_n^{-1}\int_{t-ph_n}^{t-qh_n}\int_0^s\frac{H_n(x) - H(x)}{(1 - H_n(x))(1 - H(x))}\mathrm{d}(H_{n1}(x)$$

$$- H_1(x))\mathrm{d}K\left(\frac{t-s}{h_n}\right)$$

$$+ h_n^{-1}\int_{t-ph_n}^{t-qh_n}\int_0^s \frac{H_n(x) - H(x)}{(1-H_n(s))(1-H(s))}\mathrm{d}H_1(x)\mathrm{d}K\left(\frac{t-s}{h_n}\right)$$

$$\stackrel{\mathrm{d}}{=} \Delta_{41n} + \Delta_{42n},$$

而由 Glivenko-Cantelli 引理,显然有

$$|\Delta_{41n}| \leqslant C(nh_n)^{-1}\log\log n. \qquad (2.2.37)$$

注意到 $\mathrm{d}H_1 = (1-G(s))f(s)\mathrm{d}s$,且通过分部积分我们有

$$\Delta_{42n} = \int_p^q K(u)\frac{H_n(t-h_nu) - H(t-h_nu)}{(1-H_n(t-h_nu))(1-H(t-h_nu))}$$
$$\times (1-G(t-h_nu))f(t-h_nu))\mathrm{d}u.$$

又由条件 $C4H$,知对任何 $\tau' > 0$,$f(t)$ 在$[0,\tau']$有界.因此根据 Glivenko-Cantelli 引理与引理 2.2.1,得到

$$|\Delta_{42n}| \leqslant Cn^{-\frac{1}{2}}\sqrt{\log\log n}. \qquad (2.2.38)$$

易见

$$|\Delta_{5n}| \leqslant \sup_{0\leqslant t\leqslant\tau}|I_{12n}(t)|h_n^{-1}\int_{t-ph_n}^{t-qh_n}|\mathrm{d}K(u)|.$$

由(2.2.9),我们有

$$|\Delta_{5n}| \leqslant C(nh_n)^{-1}. \qquad (2.2.39)$$

注意到

$$\begin{cases}\tilde{f}_n(t) = (nh_n)^{-1}\sum_{i=1}^n K\left(\frac{t-X_i}{h_n}\right)I[\delta_i = 1],\\ \mathrm{d}H_1(t) = (1-G(t))f(t)\mathrm{d}t,\end{cases}$$

我们有

$$E\tilde{f}_n(t)$$
$$= (nh_n)^{-1}\sum_{i=1}^n\int_{t-ph_n}^{t-qh_n}E(I(\delta_i = 1|X_i = x))K\left(\frac{t-x}{h_n}\right)h(x)\mathrm{d}x$$

$$= (nh_n)^{-1}\sum_{i=1}^{n}\int_{t-ph_n}^{t-qh_n}\frac{(1-G(x))f(x)}{h(x)}K\left(\frac{t-x}{h_n}\right)h(x)\mathrm{d}x$$

$$= h_n^{-1}\int_{t-ph_n}^{t-qh_n}K\left(\frac{t-x}{h_n}\right)\mathrm{d}H_1(x). \tag{2.2.40}$$

使用(2.2.40)并分部积分,得

$$\begin{aligned}
&-\Delta_{1n}(t)\\
&=\frac{\tilde{f}_n(t)-E\tilde{f}_n(t)}{1-H(t)}+h_n^{-1}\int_{t-ph_n}^{t-qh_n}\frac{H_{n1}(s)-H_1(s)}{1-H(t)}\mathrm{d}K\left(\frac{t-s}{h_n}\right)\\
&\quad-h_n^{-1}\int_{t-ph_n}^{t-qh_n}\frac{H_{n1}(s)-H_1(s)}{1-H(s)}\mathrm{d}K\left(\frac{t-s}{h_n}\right)\\
&=\frac{\tilde{f}_n(t)-E\tilde{f}_n(t)}{1-H(t)}\\
&\quad+h_n^{-1}\int_{t-ph_n}^{t-qh_n}\frac{(H_{n1}(s)-H_1(s))(H(t)-H(s))}{(1-H(t))(1-H(s))}\mathrm{d}K\left(\frac{t-s}{h_n}\right).
\end{aligned}\tag{2.2.41}$$

以 $E_n(t)$ 定义(2.2.40)中最后一式的第二项,应用条件 $C4H$,及关于 $H(t)-H(s)$ 的微分中值定理,得

$$\begin{aligned}
E_n(t)&\leqslant C\sup_{0\leqslant t\leqslant\tau}|H_{n1}(s)-H_1(s)|(1-H(\tau))^{-2}\\
&\quad\times\left|\int_{t-ph_n}^{t-qh_n}\frac{|t-s|}{h_n}\mathrm{d}K\left(\frac{t-s}{h_n}\right)\right|\\
&\leqslant Cn^{-\frac{1}{2}}\sqrt{\log\log n},\quad a.s.
\end{aligned}\tag{2.2.42}$$

因此(2.2.34),(2.2.35),(2.2.36),(2.2.37),(2.2.38),(2.2.39),(2.2.41)与(2.2.42)一起证明了定理 2.2.4.

**定理 2.2.5** 在定理 2.2.4 的假定下,若 $h_n\log\log n\to 0$,且 $(nh_n)^{-\frac{1}{2}}\log\log n\to 0$,对任何 $0<t<\tau$,我们有

$$(nh_n)^{\frac{1}{2}}(\lambda_n(t)-\tilde{\lambda}_n(t))\to N\left(0,\lambda(t)\int_p^q K^2(u)\mathrm{d}u\right).$$

证明:由定理 2.2.4 及条件

$$h_n \log\log n \to 0, \quad (nh_n)^{-\frac{1}{2}} \log\log n \to 0,$$

我们仅需证

$$(nh_n)^{\frac{1}{2}}(\tilde{f}_n(t) - E\tilde{f}_n(t)) \to N\left(0, (1 - G(t))f(t)\int_p^q K^2(u)\mathrm{d}u\right). \tag{2.2.43}$$

设 $Y_i = (X_i, \delta_i)$，且

$$W_n(Y_i) = h_n^{-1} K\left(\frac{t - x_i}{h_n}\right) I[\delta_i = 1].$$

因此，$\tilde{f}_n(t)$ 能被写成独立且与

$$W_n = h_n^{-1} K\left(\frac{t - x}{h_n}\right) I[\delta = 1]$$

同分布的随机变量的平均，即

$$\tilde{f}_n(t) = n^{-1} \sum_i W_n(X_i). \tag{2.2.44}$$

对某 $\delta > 0$

$$\begin{aligned}
&E|W_n|^{2+\delta} \\
&= \int_{t-ph_n}^{t-qh_n} E(I[\delta = 1] \mid X = x)\left[h_n^{-1} K\left(\frac{t - s}{h_n}\right)\right]^{2+\delta} h(x)\mathrm{d}x \\
&= h_n^{-(2+\delta)} \int_{t-ph_n}^{t-qh_n} K^{2+\delta}\left(\frac{t - y}{h_n}\right)(1 - G(y))f(y)\mathrm{d}y \\
&= h_n^{-(1+\delta)} \int_p^q K^{2+\delta}(u)(1 - G(t - h_n u))f(t - h_n u)\mathrm{d}u \\
&= h_n^{-(1+\delta)}(1 - G(t))f(t)\int_p^q K^{2+\delta}(u)\mathrm{d}u + h_n^{-(1+\delta)} o(1).
\end{aligned} \tag{2.2.45}$$

类似地，可计算

$$\begin{aligned}
\operatorname{Var} W_n &= EW_n^2 - (EW_n)^2 \\
&= h_n^{-1}(1 - G(t))f(t)\int_p^q K^2(u)\mathrm{d}u \\
&\quad + h_n^{-1} o(1) + O(1).
\end{aligned} \tag{2.2.46}$$

既然条件$(nh_n)^{-1}\log\log n \to 0$蕴涵了$nh_n \to \infty$,根据(2.2.45),(2.2.46)我们有

$$\frac{E|W_n - EW_n|^{2+\delta}}{n^{\frac{\delta}{2}}\mathrm{Var}^{\frac{2+\delta}{2}}(W_n)} \leqslant \frac{2^{1+\delta}h_n^{1+\delta}(E|W_n|^{2+\delta} + |EW_n|^{2+\delta})}{(nh_n)^{\frac{\delta}{2}}h_n^{\frac{2+\delta}{2}}\mathrm{Var}^{\frac{2+\delta}{2}}(W_n)} \to 0. \tag{2.2.47}$$

由 Lyapounov 定理,(2.2.47)是

$$\frac{\tilde{f}_n(t) - E\tilde{f}_n(t)}{\sqrt{\mathrm{Var}\tilde{f}_n(t)}} \to N(0,1) \tag{2.2.48}$$

成立的充分条件.

由(2.2.44)与(2.2.46),可通过计算得到

$$\mathrm{Var}\tilde{f}_n(t)$$

$$= (nh_n)^{-1}\left[(1 - G(t))f(t)\int_p^q K^2(u)\mathrm{d}u + o(1) + O(h_n)\right]. \tag{2.2.49}$$

(2.2.48),(2.2.49)一起证明了(2.2.43),因此证明了定理2.2.5.

**定理 2.2.6** 在条件 $C2\lambda$ 与定理 2.2.5 条件下,若$(nh_n)^{-\frac{1}{2}}\log\log n \to 0$且$nh_n^{2k+1} \to 0$,则对任何$0 < t < \tau$我们有

$$(nh_n)^{\frac{1}{2}}(\lambda_n(t) - \lambda(t)) \to N\left(0, \lambda(t)\int_p^q K^2(u)\mathrm{d}u\right).$$

证明:由(2.2.19),(2.2.22) (2.2.30),我们有

$$(nh_n)^{\frac{1}{2}}(\lambda_n(t) - \lambda(t))$$

$$= (nh_n)^{\frac{1}{2}}(\lambda_n(t) - \tilde{\lambda}_n(t)) + O(n^{\frac{1}{2}}h_n^{k+\frac{1}{2}}). \tag{2.2.50}$$

注意到条件$nh_n^{2k+1} \to 0$蕴涵了$h_n\log\log n \to 0$,因此,定理2.2.5中的条件全满足. 于是,使用定理2.2.5与(2.2.50),我们有

$$(nh_n)^{\frac{1}{2}}(\lambda_n(t) - \lambda(t)) \to N\left(0, \lambda(t)\int_p^q K^2(u)\mathrm{d}u\right).$$

这就证明了定理2.2.6.

# §2.3 均值型泛函估计的一些渐近结果

## 2.3.1 记号与主要结果

设 $X_1,X_2,\cdots,X_n$ 是非负独立同分布表示生存时间的随机变量,有共同的分布函数 $F$, $Y_1,Y_2,\cdots,Y_n$ 是非负独立同分布表示删失的随机变量,有共同的分布函数 $G$,且假定诸 $X_i$ 独立于诸 $Y_i$. 在随机删失模型中,我们不能完全观察真的生存时间 $X_1,X_2,\cdots,X_n$,而仅能观察到

$$Z_i = \min(X_i, Y_i), \quad \delta_i = I[X_i \leqslant Y_i], \quad i = 1,2,\cdots,n,$$

其中 $I[\cdot]$ 定义某事件的示性函数.

本节考虑用上面所获的删失观察估计泛函

$$\xi_u(F) = \int_0^{\infty}(1 - F(t))\mathrm{d}\theta_u(t),$$

其中 $\theta_u(t)$ 是非负可测且使 $\xi_u(F) < \infty$ 的可能依赖某实参数 $u$ 的实值函数. 下面是 $\xi_u(F)$ 所代表的一些例子:(i) 若 $\theta_u(t) = t$, $\xi_u(F)$ 表示 $F(t)$ 的平均生存时间;(ii) 对某实可测连续不减的函数 $q(\cdot)$ 及 $u \geqslant 0$,若 $q(0) = 0$, $\theta_u(t) = I[q(t) > u]$,则 $\xi_u(F) = P(q(X) > u)$,特别地,若 $q(t) = t$,则 $\xi_u(F) = 1 - F(u)$;(iii) 若对 $u \geqslant 0$, $\theta_u(t) = \dfrac{I[t \leqslant u]}{1 - F(t^-)}$, 则当 $F(x)$ 连续时 $\xi_u(F)\left(= \displaystyle\int_0^u \frac{1}{1 - F(t)}\mathrm{d}F(t)\right)$是累积失效率函数.

下面,我们定义

$$\xi_u(\hat{F}_n) = \int_0^{Z_{(n)}}(1 - \hat{F}_n(t))\mathrm{d}\theta_u(t)$$

估计 $\xi_u(F)$, 其中 $Z_{(n)} = \max\limits_{1\leqslant i\leqslant n} Z_i$, $1 - \hat{F}_n$ 是(0.1)所定义的 Kaplan-Meier 估计.

为简便计,以下用 $\theta(t)$, $\xi(F)$ 与 $\xi(\hat{F}_n)$, 分别表示 $\theta_u(t)$,

$\xi_u(F)$ 与 $\xi_u(\hat{F}_n)$.

易见 $\theta_u(t) = t$ 时，$\xi(\hat{F}_n)$ 就是 Gill[39] 所定义的平均生存时间估计 $\hat{\mu}_n$. 通过证明 $\sqrt{n}(\hat{F}(t) - F(t))/(1 - F(t))$ 在整个半直线上弱收敛，Gill[39] 证明了 $\hat{\mu}_n$ 的渐近正态性，值得指出的是其证明是不严格的. 本文通过用点过程鞅方法证明了 $\xi(\hat{F}_n)$ 的渐近正态性，并建立了它的均方误差与一个概率不等式，而不需 $F, G$ 连续的假设. 易见 Gill[39] 的结果是本文下面定理 2.3.1 的特殊情形.

在陈述结果之前，我们需要下面记号和定义. 注意到$(Z_i, \delta_i)$, $i = 1, 2, \cdots, n$ 独立同分布且对 $0 \leqslant t < \infty$,

$$1 - H(t) = P(Z > t) = (1 - F(t))(1 - G(t)),$$

$$H_1(t) = P(Z \leqslant t, \delta = 1) = \int_{[0,t]} (1 - G(s^-))\mathrm{d}F(s),$$

$$H_0(t) = P(Z \leqslant t, \delta = 0) = \int_{[0,t]} (1 - F(s))\mathrm{d}G(s),$$

$H = H_1 + H_0$. 对任何分布函数 $K(t)$，我们定义 $\tau_K = \inf\{t : K(t) = 1\}$. 本节假定 $\tau_F \leqslant \tau_G$. 易见，$\tau_H = \tau_F \wedge \tau_G = \tau_F$. 记

$$H_n(t) = \frac{1}{n}\sum_{i=1}^n I[Z_i \leqslant t],$$

$$H_{n1}(t) = \frac{1}{n}\sum_{i=1}^n I[Z_i \leqslant t, \delta_i = 1],$$

$$H_{n0}(t) = \frac{1}{n}\sum_{i=1}^n I[Z_i \leqslant t, \delta_i = 0],$$

且设

$$\Lambda(t) = \int_{[0,t]} \frac{1}{1 - F(s^-)}\mathrm{d}F(s),$$

对任何右连续的函数 $A(t)$，记 $\Delta A(t) = A(t) - A(t^-)$ 以下约定 $\alpha$ 表示任何所需的常数，对 $s > 0$，$\int_s^t = \int_{(s,t]}$，而$\int_0^t = \int_{[0,t]}$.

**定理 2.3.1** 若条件

(i) $\int_0^{\tau_H} \frac{1}{1 - G(s^-)} \mathrm{d}F(s) < \infty$;

(ii) $\sup_t \left| \frac{\int_t^{\tau_H} (1 - F(s)) \mathrm{d}\theta(s)}{1 - F(t)} \right| < \infty$

满足,则当$\sqrt{n}\int_{Z_{(n)}}^{\tau_H} (1 - F(t)) \mathrm{d}\theta(t) \xrightarrow{p} 0$,有

$$\sqrt{n}(\xi(\hat{F}_n) - \xi(F)) \xrightarrow{\mathcal{L}} N(0, \sigma^2),$$

其中

$$\sigma^2 = \int_0^{\tau_H} \left( \int_s^{\tau_H} (1 - F(l)) \mathrm{d}\theta(l) \right)^2 \frac{1 - F(s^-)}{1 - F(s)} \frac{1}{1 - H(s^-)} \mathrm{d}\Lambda(s).$$

**定理 2.3.2** 在定理2.3.1的条件(i)与(ii)下,若存在常数序列 $m_n \uparrow \tau_H$ 使得$\liminf_n n(1 - H(m_n)) \geqslant 2\log n$ 且$\limsup_n \sqrt{n}\int_{m_n}^{\tau_H}(1 - F(l))\mathrm{d}\theta(l) < \infty$,则

$$E(\xi(\hat{F}_n) - \xi(F))^2$$

$$\leqslant 10n^{-1}\left[ \sup_s \left( \frac{\int_s^{\tau_H}(1 - F(x))\mathrm{d}\theta(x)}{1 - F(s)} \right)^2 \int_0^{\tau_H} \frac{1}{1 - G(s^-)} \mathrm{d}F(s) \right.$$

$$+ \left( \int_0^{\tau_H} (1 - F(s)) \mathrm{d}\theta(s) \right)^2$$

$$\left. + \left( \limsup_n \sqrt{n} \int_{m_n}^{\tau_H} (1 - F(l)) \mathrm{d}\theta(l) \right)^2 \right]$$

对充分大的 $n$ 成立.

**定理 2.3.3** 在定理2.3.1的条件(i)与(ii)下,对 $\forall \varepsilon > 0$,若存在常数序列 $t_\varepsilon > 0$ 使得$\int_{t_\varepsilon}^{\tau_H}(1 - F(s))\mathrm{d}F(s) \leqslant \varepsilon/8$,则当 $n$ 充分大时,有

$$P(|\xi(\hat{F}_n) - \xi(F)| > \varepsilon)$$

$$\leqslant \alpha \left[ \exp\left\{ -\frac{n\varepsilon^2}{8\eta_1} \psi\left( \frac{\varepsilon c(t_\varepsilon) n}{2\eta_1} \right) \right\} \right.$$

$$+\exp\{-\alpha(1-H(t_\varepsilon))^2(1-F(t_\varepsilon))n\}].$$

其中

$$\eta_1 = 4\sup_s\left(\frac{\int_s^{\tau_H}(1-F(t))\mathrm{d}\theta(t)}{1-F(s)}\right)^2\int_0^\infty\frac{1}{1-G(s^-)}\mathrm{d}F(s),$$

$$\psi(x)=2p(1+x)/x^2, p(x)=x(\log x-1)+1,$$

$$c(t_\varepsilon)=4\sup_s\left|\frac{\int_s^{\tau_H}(1-F(s))\mathrm{d}\theta(s)}{1-F(s)}\right|/(1-G(t_\varepsilon^-)).$$

### 2.3.2 定理的证明

记

$$\mathbf{M}_n=\sqrt{n}(H_{n1}(t)-\int_0^t(1-H_n(s^-))\mathrm{d}\Lambda(s)).\quad(2.3.1)$$

由 Shorack 与 Wellner[72] 知 $\mathbf{M}_n$ 是$[0,\infty)$上的关于 $\sigma$ 代数流

$$\mathscr{F}_t=\sigma\{Z_iI[Z_i\leqslant s],\delta_iI[Z_i\leqslant s],1\leqslant i\leqslant n,s\leqslant t\}$$

的零均值平方可积鞅,且 $\mathbf{M}_n$ 的可料变差为

$$\langle\mathbf{M}_n\rangle=\int_0^t(1-H_n(s^-))(1-\Delta\Lambda(s))\mathrm{d}\Lambda(s).\quad(2.3.2)$$

定理 2.3.1 的证明:设

$$J_n(t)=I[0\leqslant t\leqslant Z_{(n)}].$$

由 Shorack 与 Wellner[72],对 $0\leqslant t<\tau_H, 0\leqslant t\leqslant Z_{(n)}$,有

$$\sqrt{n}(\hat{F}_n(t)-F(t))$$
$$=(1-F(t))\int_0^t\frac{1-\hat{F}_n(s^-)}{1-F(s)}\frac{J_n(s)}{1-H_n(s^-)}\mathrm{d}\mathbf{M}_n(s).\quad(2.3.3)$$

设

$$\tilde{\xi}_n(F)=\int_0^{Z_{(n)}}(1-F(t))\mathrm{d}\theta(t).\quad(2.3.4)$$

因此,由(2.3.3) 并通过分部积分,可得

$$\sqrt{n}(\xi(\hat{F}_n)-\tilde{\xi}_n(F))$$

$$= \int_0^{Z_{(n)}} \left( \int_0^s \frac{1-\hat{F}_n(x^-)}{1-F(x)} \frac{J_n(x)}{1-H_n(x^-)} \mathrm{d}\mathbf{M}_n(x) \right) \mathrm{d}\int_s^{\tau_H} (1-F(x))\mathrm{d}\theta(x)$$

$$= \left( \int_0^{Z_{(n)}} \frac{1-\hat{F}_n(s^-)}{1-F(s)} \frac{J_n(s)}{1-H_n(s^-)} \mathrm{d}\mathbf{M}_n(s) \right) \int_{Z_{(n)}}^{\tau_H} (1-F(x))\mathrm{d}\theta(x)$$

$$- \int_0^{Z_{(n)}} \left( \int_s^{\tau_H} (1-F(x))\mathrm{d}\theta(x) \right) \frac{1-\hat{F}_n(s^-)}{1-F(s)} \frac{J_n(s)}{1-H_n(s^-)} \mathrm{d}\mathbf{M}_n(s)$$

$$= - \int_0^{Z_{(n)}} \left[ \int_s^{Z_{(n)}} (1-F(x))\mathrm{d}\theta(x) \right] \frac{1-\hat{F}_n(s^-)}{1-F(s)} \frac{J_n(s)}{1-H_n(s^-)} \mathrm{d}\mathbf{M}_n(s). \tag{2.3.5}$$

又令

$$h_n(t) = \left( \int_t^{Z_{(n)}} (1-F(x))\mathrm{d}\theta(x) \right) \frac{1-\hat{F}_n(t^-)}{1-F(t)} \frac{J_n(t)}{1-H_n(t^-)},$$

$$0 \leqslant t < \tau_H.$$

$$\widetilde{\mathbf{M}}_n(t) = - \int_0^t h_n(s)\mathrm{d}\mathbf{M}_n(s), \quad 0 \leqslant t < \tau_H. \tag{2.3.6}$$

易见 $h_n(s)$ 是可料过程，由于 $\mathbf{M}_n(t)$ 是$[0,\tau_H)$ 上的零均值平方可积鞅且由条件

$$\sup_s \left| \int_s^{\tau_H} (1-F(x))\mathrm{d}\theta(x)/(1-F(s)) \right| < \infty$$

得

$$E\int_0^t (h_n(s))^2 \mathrm{d}\langle \mathbf{M}_n \rangle(s)$$

$$\leqslant \int_0^{\tau_H} \left( \frac{\int_s^{\tau_H} (1-F(x))\mathrm{d}\theta(x)}{1-F(s)} \right)^2 E \frac{J_n(s)}{1-H_n(s^-)} \mathrm{d}\Lambda(s)$$

$$\leqslant \sup_s \left( \frac{\int_s^{\tau_H} (1-F(x))\mathrm{d}\theta(x)}{1-F(s)} \right)^2$$

$$\times \int_0^{\tau_H} \sum_{k=1}^n \frac{1}{k} \binom{n}{k} (1-H(s^-))^k H^{n-k}(s^-) \frac{\mathrm{d}F}{1-F(s^-)}$$

$$\leqslant \sup_s \left\{ \frac{\int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x)}{1-F(s)} \right\}^2$$

$$\times \sum_{k=1}^{n} \frac{1}{k}\binom{n}{k}\int_0^{\tau_H}(1-G(s^-))\mathrm{d}F(s) < \infty,$$

因而由 Shorack 与 Wellner[72] 中附录 B 的定理 3.1,知 $\mathbf{M}_n(t)$ 是$[0, \tau_H)$ 上的零均值平方可积鞅,因而,我们有

$$\langle \widetilde{\mathbf{M}}_n \rangle(t) = \int_0^t h_n^2(s)(1-H_n(s^-))(1-\Delta\Lambda(s))\mathrm{d}\Lambda(s), 0 \leqslant t \leqslant \tau_H. \tag{2.3.7}$$

由 Shorack 与 Wellner[72] 中第七章定理 3.1 和 Glivenko-Ct-phlli 的定理,及定理 2.3.1 中的条件(i),(ii) 知

$$\langle \widetilde{\mathbf{M}}_n \rangle(t) \xrightarrow{a.s.} \int_0^t \left( \int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x) \right)^2$$

$$\times \left( \frac{1-F(s^-)}{1-F(s)} \right)^2 \frac{(1-\Delta\Lambda(s))}{1-H(s^-)} \frac{\mathrm{d}F(s)}{1-F(s^-)}$$

$$\leqslant \int_0^t \left\{ \frac{\int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x)}{1-F(s)} \right\}^2 \frac{\mathrm{d}F(s)}{1-G(s^-)}$$

$$\leqslant \sup_s \left\{ \frac{\int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x)}{1-F(s)} \right\}^2 \int_0^t \frac{\mathrm{d}F(s)}{1-G(s^-)} < \infty, \tag{2.3.8}$$

设

$$M_{1i}(t) = I[Z_i \leqslant t, \delta_i = 1] - \int_0^t I[Z_i \geqslant s]\mathrm{d}\Lambda(s).$$

易见

$$\mathbf{M}_n(t) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} M_{1i}(t). \tag{2.3.9}$$

将(2.3.9) 代入(2.3.6),我们有

$$\widetilde{\mathbf{M}}_n(t) = -\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_0^t h_n(s)\mathrm{d}M_{1i}(s). \qquad (2.3.10)$$

为了证明 $\widetilde{\mathbf{M}}_n(t)$ 的渐近正态性，除(2.3.8)外，还需验证 Linderberge 条件，也就是要验证

$$\sum_{i=1}^{n}\int_0^t B_n(s)I[B_n(s) > \varepsilon]\mathrm{d}\langle M_{1i}\rangle(s) \xrightarrow{p} 0, \quad (2.3.11)$$

对任何 $\varepsilon > 0$ 成立，其中 $B_n(s) = \left(\frac{1}{\sqrt{n}}h_n(s)\right)^2$.

由 Shorack 与 Wellner[72] 中第7章定理3.1，既然 $\hat{F}_n(s^-) \xrightarrow{a.s.} F(s^-), H_n(s^-) \xrightarrow{a.s.} H(s^-)$ 在 $s \in [0,t]$ 上一致成立，因而再由 Glivenko-Cantelli 定理，我们有 $B_n(s) \xrightarrow{a.s.} 0$ 在 $s \in [0,t]$ 上一致地成立. 因此对任何 $\varepsilon > 0$ 当 $n$ 充分大时有 $I[B_n(s) > \varepsilon] = 0, a.s.$ 在 $s \in [0,t]$ 上一致地成立. 由此推得(2.3.11)左边以概率1为0，于是(2.3.11)得证. 而由(2.3.8)，(2.3.11)和鞅的中心极限定理，我们得到

$$\widetilde{\mathbf{M}}_n(t) \xrightarrow{\mathcal{L}} N(0,\sigma^2(t)), 0 \leqslant t < \tau_H, \qquad (2.3.12)$$

其中

$$\sigma^2(t) = \int_0^t\left(\int_s^{\tau_H}(1 - F(x))\mathrm{d}\theta(x)\right)^2\left[\frac{1 - F(s^-)}{1 - F(s)}\right]\frac{1}{1 - H(s^-)}\mathrm{d}\Lambda(s). \qquad (2.3.13)$$

由(2.3.5)，(2.3.6)，易见

$$\begin{aligned}\widetilde{\widetilde{\mathbf{M}}}_n(t) &\stackrel{\mathrm{d}}{=} \sqrt{n}((\hat{F}_n) - \tilde{\xi}_n(F)) - \widetilde{\mathbf{M}}_n(t) \\ &= -\int_t^{Z_{(n)}} h_n(s)\mathrm{d}\mathbf{M}_n(s). \end{aligned} \qquad (2.3.14)$$

注意到

$$\mathbf{M}_n(s) = \sqrt{n}(H_{n1}(s) - H_1(s)) + \int_0^s \sqrt{n}\,\frac{H_n(x^-) - H(x^-)}{H(x^-)}\mathrm{d}H_1(x), \qquad (2.3.15)$$

因此

$$\tilde{\tilde{\mathbf{M}}}_n(t) = -\sqrt{n}\int_t^{Z_{(n)}} h_n(s)\mathrm{d}(H_{n1}(s) - H_1(s))$$

$$- \int_t^{Z_{(n)}} h_n(s)\frac{\sqrt{n}(H_n(s^-) - H(s^-))}{H(s^-)}\mathrm{d}H_1(s)$$

$$= \tilde{\tilde{\mathbf{M}}}_{n1}(t) + \tilde{\tilde{\mathbf{M}}}_{n2}(t). \tag{2.3.16}$$

由 Major 与 Rejtö[54] 知,存在 Brown 桥 $B(t), 0 \leqslant t \leqslant 1$ 使得

$$\sup_{0 \leqslant s \leqslant \tau_H} |\sqrt{n}(H_{n1}(s) - H_1(s)) - B(H_1(s))| \xrightarrow{a.s.} 0. \tag{2.3.17}$$

设 $h(t) = \left(\int_t^{\tau_H}(1 - F(x))\mathrm{d}\theta(x)\right)\frac{1 - F(t^-)}{1 - F(t)}\frac{1}{1 - H(t^-)}$. 反复分部积分,由(2.3.17),Shorack 与 Wellner[72] 中第七章定理 3.1 和切比雪夫不等式,可得对任何 $\varepsilon > 0$,当 $t \to \tau_H$ 时

$$P(|\tilde{\tilde{\mathbf{M}}}_{n1}(t)| > \varepsilon)$$

$$\leqslant P(2\sup_{0 \leqslant s \leqslant \tau_H} |\sqrt{n}(H_{n1}(s) - H_1(s)) - B(H_1(s))| h_n(t) > \varepsilon/2)$$

$$+ P(|\int_t^{\tau_H} h_n(s)\mathrm{d}B(H_1(s))| > \varepsilon/2)$$

$$\to P(|\int_t^{\tau_H} h(s)\mathrm{d}B(H_1(s))| > \varepsilon/2)$$

$$\leqslant \frac{2}{\varepsilon}E|\int_t^{\tau_H} h(s)\mathrm{d}B(H_1(s))| \leqslant \frac{2}{\varepsilon}\left(\int_t^{\tau_H} h^2(s)\mathrm{d}H_1(s)\right)^{\frac{1}{2}}$$

$$\leqslant \frac{2}{\varepsilon}\sup_s\left|\frac{\int_s^{\tau_H}(1 - F(s))\mathrm{d}\theta(s)}{1 - F(s)}\right|\left(\int_t^{\tau_H}\frac{1}{1 - G(s^-)}\mathrm{d}F(s)\right)^{\frac{1}{2}} \to 0, \tag{2.3.18}$$

类似可证当 $n \to \infty, t \to \tau_H$ 时

$$P(|\tilde{\tilde{\mathbf{M}}}_{n2}(t) > \varepsilon) \to 0, \tag{2.3.19}$$

而由(2.3.14),(2.3.12),(2.3.16),(2.3.18) 与(2.3.19),得到当 $n \to \infty$,并取 $t$ 充分接近 $\tau_H$ 时

$$\sqrt{n}(\xi(\hat{F}_n) - \tilde{\xi}_n(F)) = \widetilde{\mathbf{M}}_n(t) + \tilde{\widetilde{\mathbf{M}}}_n(t) \xrightarrow{\mathcal{L}} N(0, \sigma^2), \tag{2.3.20}$$

其中 $\sigma^2 = \sigma^2(\tau_H)$,$\sigma^2(\cdot)$ 如(2.3.13) 所定义,易见

$$\begin{aligned}&\sqrt{n}(\xi(\hat{F}_n) - \xi(F))\\ &= \sqrt{n}(\xi(\hat{F}_n) - \tilde{\xi}_n(F)) - \sqrt{n}\int_{Z_{(n)}}^{\tau_H}(1 - F(t))\mathrm{d}\theta(t).\end{aligned} \tag{2.3.21}$$

(2.3.20),(2.3.21) 与条件$\sqrt{n}\int_{Z_{(n)}}^{\tau_F}(1 - F(t))\mathrm{d}\theta(t) \xrightarrow{p} 0$ 一起,即给出

$$\sqrt{n}(\xi(\hat{F}_n) - \xi(F)) \xrightarrow{\mathcal{L}} N(0, \sigma^2).$$

此处 $\sigma^2$ 如前所定义. 于是定理 2.3.1 得证

定理2.3.2 的证明:注意到$\sqrt{n}(\xi(\hat{F}_n) - \tilde{\xi}_n(F))$是(2.3.6) 所定义的鞅 $\widetilde{\mathbf{M}}_n$ 在 $t = Z_{(n)}$ 处的赋值,于是有

$$\begin{aligned}&nE(\xi(\hat{F}_n) - \tilde{\xi}_n(F))^2 = E(\langle\widetilde{\mathbf{M}}_n\rangle(t))\big|_{t = Z_{(n)}}\\ &= E\int_0^{Z_{(n)}}\left(\int_s^{Z_{(n)}}(1 - F(x))\mathrm{d}\theta(x)\right)^2\left(\frac{1 - \hat{F}_n(s^-)}{1 - F(s)}\right)^2\\ &\quad\times\frac{J_n}{1 - H_n(s^-)}(1 - \Delta\Lambda(s))\mathrm{d}\Lambda(s)\\ &\leqslant \int_0^{\tau_H}E^{\frac{1}{2}}\left[\left(\int_s^{\tau_H}(1 - F(x))\mathrm{d}\theta(x)\right)\frac{1 - \hat{F}_n(s^-)}{1 - F(s)}J_n\right]^4\\ &\quad\times E^{\frac{1}{2}}\frac{J_n}{(1 - H_n(s^-))^2}\mathrm{d}\Lambda(s).\end{aligned} \tag{2.3.23}$$

因

$$\left[\left(\int_s^{\tau_H}(1 - F(x))\mathrm{d}\theta(x)\right)\frac{1 - \hat{F}_n(s^-)}{1 - F(s)}J_n\right]^4$$

$$\leqslant \sup_s \left[\frac{\int_s^{\tau_H}(1-F(x))\mathrm{d}x}{1-F(s)}\right]^4 < \infty,$$

于是由有界控制收敛定理和 Shorack 与 Wellner[72] 第七章中的定理 3.1,知

$$E\left[\left(\int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x)\right)\left(\frac{1-\hat{F}_n(s^-)}{1-F(s)}\right)J_n\right]^4$$

$$\to\left[\left(\int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x)\right)\left(\frac{1-F(s^-)}{1-F(s)}\right)\right]^4 \qquad (2.3.24)$$

在 $s\in[0,\tau_H]$ 上一致地成立. 因此,当 $n$ 充分大时,我们有

$$E^{\frac{1}{2}}\left[\left(\int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x)\right)\left(\frac{1-\hat{F}_n(s^-)}{1-F(s)}\right)J_n\right]^4$$

$$\leqslant 2\left[\int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x)\,\frac{1-F(s^-)}{1-F(s)}\right]^2 \qquad (2.3.25)$$

在 $s\in[0,\tau_H]$ 上一致地成立. 又

$$E\left(\frac{J_n}{1-H_n(s^-)}\right)^2$$

$$= n^2\sum_{k=1}^{n}\frac{1}{k^2}\binom{n}{k}(1-H(s^-))^k H^{n-k}(s^-)$$

$$\leqslant \sum_{k=1}^{n}\left|\frac{n^2}{k^2}-\frac{1}{(1-H(s^-))^2}\right|\binom{n}{k}(1-H(s^-))^k H^{n-k}(s^-)$$

$$+\frac{1}{(1-H(s^-))^2}(1-H^n(s^-)). \qquad (2.3.26)$$

而对任何 $\varepsilon>0$,存在 $\delta>0$ 使得当 $\sup_s\left|\frac{k}{n}-(1-H(s^-))\right|<\delta$ 时,有 $\sup_s\left|\frac{n^2}{k^2}-\frac{1}{(1-H(s^-))^2}\right|<\varepsilon$. 令 $A_{n\delta}=\left\{k:\sup_s\left|\frac{k}{n}-(1-H(s^-))\right|<\delta\right\}$,$A_{n\delta}^c$ 是 $A_{n\delta}$ 的补集. 因此,由(2.3.26)我们有

$$E\left(\frac{J_n}{1-H_n(s^-)}\right)^2$$

$$\leqslant \sum_{A_{n\delta}} \sup_s \left\{ \left| \frac{n^2}{k^2} - \frac{1}{(1-(s^-))^2} \right| \right\} \binom{n}{k} (1-H(s^-))^k H^{n-k}(s^-)$$

$$+ \sum_{A_{n\delta}^c} \sup_s \left\{ \left| \frac{n^2}{k^2} - \frac{1}{(1-H(s^-))^2} \right| \right\} \binom{n}{k} (1-H(s^-))^k H^{n-k}(s^-)$$

$$+ \frac{1}{(1-H(s^-))^2}(1-H^n(s^-))$$

$$\leqslant \varepsilon + \left( n^2 + \frac{1}{(1-H(s^-))^2} \right) P(\sup_s |H_n(s^-) - H(s^-)| \geqslant \delta)$$

$$+ \frac{1}{(1-H(s^-))^2}(1-H^n(s^-))$$

$$\leqslant \varepsilon + \left( n^2 + \frac{1}{(1-H(s^-))^2} \right) e^{-2n\delta^2} + \frac{1}{(1-H(s^-))^2}(1-H^n(s^-)). \tag{2.3.27}$$

由(2.3.27),我们得

$$E^{\frac{1}{2}} \left( \frac{J_n}{1-H_n(s^-)} \right)^2 \leqslant \frac{1}{1-H(s^-)} + o(1) \tag{2.3.28}$$

在 $s \in [0, \tau_H]$ 上一致地成立. 于是(2.3.23),(2.3.25)与(2.3.28)一起可得

$$nE(\xi(\hat{F}_n) - \tilde{\xi}_n(F))^2$$

$$\leqslant 4\int_0^{\tau_H} \left( \int_s^{\tau_H} (1-F(x))\mathrm{d}\theta(x) \frac{1-F(s^-)}{1-F(s)} \right)^2 \frac{1}{1-H(s^-)} \frac{\mathrm{d}F(s)}{1-F(s^-)}$$

$$+ o(1)\int_0^{\tau_H} \left( \int_s^{\tau_H} (1-F(x))\mathrm{d}\theta(x) \frac{1-F(s^-)}{1-F(s)} \right)^2 \frac{\mathrm{d}F(s)}{1-F(s^-)}$$

$$\leqslant 5 \sup_s \left[ \frac{\int_s^{\tau_H} (1-F(x)\mathrm{d}\theta(x)}{1-F(s)} \right]^2 \int_0^{\tau_H} \frac{1}{1-G(s^-)} \mathrm{d}F(s), \tag{2.3.29}$$

显然,对充分大的 $m_n$,我们有

$$nE(\xi(\hat{F}_n) - \xi(F))^2$$

$$\leqslant 2nE(\xi(\hat{F}_n) - \tilde{\xi}_n(F))^2 + 2n\left( \int_{Z_{(n)}}^{\tau_H} (1-F(t))\mathrm{d}\theta(t) \right)^2$$

$$
\begin{aligned}
&\leqslant 2nE(\xi(\hat{F}_n)-\tilde{\xi}_n(F))^2 \\
&\quad + 2nE\left[\left(\int_{Z_{(n)}}^{\tau_H}(1-F(t))\mathrm{d}\theta(t)\right)^2 I[Z_{(n)}\geqslant m_n]\right] \\
&\quad + 2nE\left[\left(\int_{Z_{(n)}}^{\tau_H}(1-F(t))\mathrm{d}\theta(t)\right)^2 I[Z_{(n)}< m_n]\right] \\
&\leqslant 2nE(\xi(\hat{F}_n)-\tilde{\xi}_n(F))^2 + 2n\left(\int_{m_n}^{\tau_H}(1-F(t))\mathrm{d}\theta(t)\right)^2 \\
&\quad + 2n\left(\int_0^{\tau_H}(1-F(t))\mathrm{d}\theta(t)\right)^2\exp\{-n(1-H(m_n))\}.
\end{aligned}
\tag{2.3.30}
$$

最后由定理 2.3.2 的条件及(2.3.29),(2.3.30) 知定理 2.3.2 得证.

定理 2.3.3 的证明:设 $\tilde{\tilde{\mathbf{M}}}_n(t)=\frac{1}{\sqrt{n}}\widetilde{\mathbf{M}}_n(t), 0\leqslant t<\tau_H$,此处 $\widetilde{\mathbf{M}}_n(t)$ 如(2.3.6) 所定义.类似于(2.3.5),我们有

$$
\begin{aligned}
\tilde{\tilde{\mathbf{M}}}_n(t) = &-\frac{1}{\sqrt{n}}\int_0^t\left[\int_0^t(1-F(x)\mathrm{d}\theta(x)\right] \\
&\times\frac{1-\hat{F}_n(s^-)}{1-F(s)}\frac{J_n(s)}{1-H_n(s^-)}\mathrm{d}\mathbf{M}_n(s).
\end{aligned}
\tag{2.3.31}
$$

易见 $\tilde{\tilde{\mathbf{M}}}_n$ 是零均值平方可积鞅,于是使用(2.3.9) 及证(2.3.7) 同样的道理,并由 $1-\hat{F}_n(s^-)\leqslant 2(1-F(s^-)), 1/(1-H_n(s))\leqslant 2/(1-H(s))$, $a.s.$ 在 $s\in[0,t]$ 上一致成立这一事实,可得

$$
\begin{aligned}
&\langle\tilde{\tilde{\mathbf{M}}}_n\rangle(t) \\
&=\frac{1}{n^2}\sum_{i=1}^n\int_0^t\left(\int_s^{\tau_{(n)}}(1-F(x))\mathrm{d}\theta(x)\right)^2\left(\frac{1-\hat{F}_n(s^-)}{1-F(s)}\right)^2\frac{J_n}{(1-H_n(s^-))^2} \\
&\quad\times I[Z_i>s](1-\Delta\Lambda(s))\mathrm{d}\Lambda(s) \\
&\overset{a.s.}{\leqslant}\frac{4}{n}\sup_s\left(\frac{\int_s^{\tau_H}(1-F(x)\mathrm{d}\theta(x)}{1-F(s)}\right)^2\int_0^t\frac{1}{1-G(s^-)}\mathrm{d}F(s),
\end{aligned}
\tag{2.3.32}
$$

类似地，我们有

$$\begin{aligned}
&\sup_{s\leqslant t}|\Delta\tilde{\bar{\mathbf{M}}}_n| \\
&\leqslant \sup_{s\leqslant t}\left[\left(\int_s^{\tau_H}(1-F(x))\mathrm{d}\theta(x)\right)\frac{1-\hat{F}_n(s^-)}{1-F(s)}\frac{J_n}{1-H_n(s^-)}\right] \\
&\overset{a.s.}{\leqslant}\frac{4\sup_s\left|\dfrac{\int_s^{\tau_H}(1-F(s))\mathrm{d}\theta(s)}{1-F(s)}\right|}{1-G(t^-)}.
\end{aligned} \tag{2.3.33}$$

设

$$\eta_n(t)=\frac{4}{n}\sup_s\left(\frac{\int_s^{\tau_H}(1-F(x)\mathrm{d}\theta(x)}{1-F(s)}\right)^2\int_0^t\frac{1}{1-G(s^-)}\mathrm{d}F(s),$$

$$c(t)=\frac{4\sup_s\left|\dfrac{\int_s^{\tau_H}(1-F(s))\mathrm{d}\theta(s)}{1-F(s)}\right|}{1-G(t^-)},$$

容易看见对任何 $\varepsilon>0$,

$$\begin{aligned}
&P(\sup_{0\leqslant s\leqslant t}\tilde{\bar{\mathbf{M}}}_n(s)|>\frac{\varepsilon}{2}) \\
&\leqslant P(\sup_{0\leqslant s\leqslant t}|\tilde{\bar{\mathbf{M}}}_n(s)|>\frac{\varepsilon}{2},\langle\tilde{\bar{\mathbf{M}}}_n\rangle(t)\leqslant\eta_n,\sup_{s\leqslant t}|\Delta\tilde{\bar{\mathbf{M}}}_n(s)|\leqslant c(t)) \\
&\quad+P(\langle\tilde{\bar{\mathbf{M}}}_n\rangle(t)>\eta_n)+P(\sup_{s\leqslant t}|\Delta\tilde{\bar{\mathbf{M}}}_n(s)|>c(t)) \\
&\overset{\mathrm{d}}{=}R_{1n}+R_{2n}+R_{3n}.
\end{aligned} \tag{2.3.34}$$

由 Shorack 与 Wellner[72] 附录中不等式 6.1，我们有

$$R_{1n}\leqslant 2\exp\left\{-\frac{n\varepsilon^2}{8\eta_1(t)}\psi\left(\frac{\varepsilon c(t)n}{2\eta_1(t)}\right)\right\}, \tag{2.3.35}$$

其中 $\psi(x),p(x)$ 如定理 2.3 所定义．设

$$E_n=\{\omega:1-\hat{F}_n(s^-)>2(1-F(s^-)),$$

$$\text{或}\frac{1}{1-H_n(s^-)}>\frac{2}{1-H(s^-)},\text{对 } s\in[0,t]\text{ 一致成立}\}.$$

注意到

$$\{\langle \overline{\widetilde{\mathbf{M}}}_n \rangle(t) > \eta_n\} \subset E_n,$$

$$\{\sup_{0 \leqslant s \leqslant t} |\Delta \overline{\widetilde{\mathbf{M}}}_n| > c(t)\} \subset E_n.$$

因此,由 Major 与 Rejtö[54],对 $i = 2,3$,我们有

$$\begin{aligned} R_{in} \leqslant & P(\sup_{0 \leqslant s \leqslant t} |\hat{F}_n(s^-) - F(s^-)| > 1 - F(s^-)) \\ & + P(\sup_{0 \leqslant s \leqslant t} |H_n(s^-) - H(s^-)| > \frac{1}{2}(1 - H(t^-)) \\ \leqslant & \alpha(\exp\{-\alpha(1 - H(t^-))^2(1 - F(t^-))n\} \\ & + \exp\left\{-\frac{(1 - H(t^-))^2}{2}n\right\} \end{aligned} \tag{2.3.36}$$

(注意 $\alpha$ 在不同的地方可表示不同的常数).

于是,(2.3.34)—(2.3.36) 一起即得 $n$ 充分大时

$$\begin{aligned} P(\sup_{0 \leqslant s \leqslant t} |\overline{\widetilde{\mathbf{M}}}_n(s)| > \varepsilon/2) \leqslant & \alpha\left[\exp\left\{-\frac{n\varepsilon^2}{8\eta_1}\psi\left(\frac{\varepsilon c(t)n}{2\eta_1}\right)\right\}\right] \\ & + \exp\{-\alpha(1 - H(t^-))^2(1 - F(t^-))n\}]. \end{aligned} \tag{2.3.37}$$

选择 $t_\varepsilon$ 充分大使得$\int_{t_\varepsilon}^{\tau_H}(1 - F(s))\mathrm{d}\theta(s) \leqslant \varepsilon/8$.易见对上述 $\varepsilon > 0$,有

$$\begin{aligned} & P(|\xi(\hat{F}_n) - \xi(F)| > \varepsilon) \\ \leqslant & P(\sup_{0 \leqslant s \leqslant t_\varepsilon} |\overline{\widetilde{\mathbf{M}}}_n(s)| > \varepsilon/2) \\ & + P\left(\int_{t_\varepsilon}^{Z_{(n)}}(1 - \hat{F}_n(s))\mathrm{d}\theta(s) > \varepsilon/4\right). \end{aligned} \tag{2.3.38}$$

用 $\Delta_n$ 定义(2.3.38) 右边第二项,并选择充分大的 $t_\varepsilon$,使得$\int_{t_\varepsilon}^{\tau_H}(1 - F(s))\mathrm{d}\theta(s) \leqslant \varepsilon/8$,于是由 Major 与 Rejtö[54] 中的(3.18),有

$$\Delta_n = P(\int_{t_\varepsilon}^{\tau_H}(1 - \hat{F}_n(s))\mathrm{d}s > \varepsilon/4, 1 - \hat{F}_n(s)$$

$\leqslant 2(1-F(s))$ 对 $s\in[0,t_\varepsilon]$ 一致成立)

$+P(1-\hat{F}_n(s)>2(1-F(s))$ 对 $s\in[0,t_\varepsilon]$ 一致成立)

$$\leqslant \alpha\exp\{-\alpha(1-H(t_\varepsilon))^2(1-F(t_\varepsilon))n\}, \tag{2.3.39}$$

最后综合(2.3.37)—(2.3.39)就得

$$P(|\xi(\hat{F}_n)-\xi(F)|>\varepsilon)\leqslant \alpha\left[\exp\left\{-\frac{n\varepsilon^2}{8\eta_1}\psi\left(\frac{\varepsilon c(t_\varepsilon)n}{2\eta_1}\right)\right\}\right.$$
$$\left.+\exp\{-\alpha(1-H(t_\varepsilon))^2(1-F(t_\varepsilon))n\}\right].$$

这就证明了定理2.3.3.

# 第3章 删失回归模型

## §3.1 非参数回归模型中加权核估计的一些收敛性质

### 3.1.1 模型与记号

考虑下面固定设计回归模型

$$Y_i = g(x_i) + e_i, \quad i = 1,2,\cdots,n. \tag{3.1.1}$$

其中 $Y_1,Y_2,\cdots,Y_n$ 是在固定设计点$x_1,x_2,\cdots,x_n$的测量观察值，$e_1,e_2,\cdots,e_n$ 是独立均值为零的随机误差序列，$g(\cdot)$ 是定义在 $[0,1]$上的未知回归函数. 不失一般性，我们假定$0 \stackrel{d}{=} x_0 < x_1 < \cdots < x_n \stackrel{d}{=} 1$.

为估计 $g(x), x \in [0,1]$，人们已提出很多估计方法（见 Berger[4] 的介绍），其中 Priesley 与 Chao[61] 提出如下加权核估计：

$$g_n(x) = \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{h_n} K\left(\frac{x - x_i}{h_n}\right) Y_i, \quad x \in [0,1]. \tag{3.1.2}$$

此处 $K(\cdot)$ 为核函数，$\{h_n\}$ 是常数序列且 $h_n \to 0$. Priesley 与 Chao[61] 在 $g(\cdot), K(\cdot)$ 分别满足 $\alpha, \beta$ 阶 Lipstchitz 条件，$e_1, e_2, \cdots, e_n$ 独立同分布，$Ee_1^2 < \infty$ 等条件下证明了 $g_n(x) \xrightarrow{p} g(x), x \in [0,1]$；而 Benedetti[3] 通过加强 $K(\cdot)$ 的限制条件，并在要求 $Ee_1^4 < \infty$ 等条件下证明了 $g_n(x) \xrightarrow{a.s.} g(x), x \in [0,1]$；王启华[104] 进一步加强 $K(\cdot)$ 的限制条件，但仅要求 $Ee_1^2 < \infty$，也证明了 Benedetti[3] 中同样的强相合性结论.

这里考虑 $Y_i, 1 \leqslant i \leqslant n$，因受到某种干扰可能被删失因而不能完全被观察的情况，即受干扰后仅能观察到

$$Z_i = \min\{Y_i, T_i\}, \quad \delta_i = I[Y_i \leqslant T_i], \quad i = 1,2,\cdots,n,$$

其中 $T_i$ 称为随机删失时间，$i = 1,2,\cdots,n$. $I[\cdot]$ 表示某事件的示性函数. 现在我们的问题是利用删失数据 $(Z_i, \delta_i), i = 1,2,\cdots,n$ 来构造 $g(x), x \in [0,1]$ 的估计.

### 3.1.2 加权核估计

考虑到这类问题的实际背景，我们假定 $Y_i: 1 \leqslant i \leqslant n, T_i: 1 \leqslant i \leqslant n$ 均为非负独立的随机变量且诸 $T_i$ 独立于诸 $Y_i$. 设 $Y_i$ 的分布函数为 $F_i$，而 $T_i$ 有共同的分布函数 $G$，且记 $F_{iS} \doteq 1 - F_i, i = 1,2,\cdots,n, \bar{F}_S = n^{-1}\sum F_{iS}, G_S = 1 - G, H_{iS} = F_{iS}G_S, \bar{H}_S = n^{-1}\sum_{i=1}^{n} H_{iS}$，对任何分布函数 $F$，定义 $\tau_F = \inf\{t: F(t) = 1\}$，$\forall p > 0, F^{-p}(t) = \left(\frac{1}{F(t)}\right)^p$.

以下我们均假定 $\tau_{F_i} \leqslant \tau_G, i = 1,2,\cdots,n$. 由于

$$E\delta_i Z_i G_S^{-1}(Z_i) = \int_0^{\tau_{F_i}}\int_y^{\tau_G} \frac{y}{1 - G(y)} \mathrm{d}G(t)\,\mathrm{d}F_i(y) = EY_i = g(x_i). \tag{3.1.3}$$

因而我们认为 $\{\delta_i Z_i G_S^{-1}(Z_i), 1 \leqslant i \leqslant n\}$ 遵从如下模型：

$$\delta_i Z_i G_S^{-1}(Z_i) = g(x_i) + \varepsilon_i, \quad i = 1,2,\cdots,n. \tag{3.1.4}$$

此处 $\varepsilon_i: 1 \leqslant i \leqslant n$ 为独立均值为零的误差序列. 于是当 $G$ 已知时,利用上面模型(3.1.4)及 Priesley 与 Chao[61] 的估计方法可得此种情况下的加权核估计为

$$g_n^{(1)}(x) = \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{h_n} K\left(\frac{x - x_i}{h_n}\right) \delta_i Z_i G_S^{-1}(Z_i), \quad x \in [0,1]. \tag{3.1.5}$$

当 $G$ 未知时,为估计 $g(x)$,我们进一步作如下假定:

**假定 3.1.1** $\forall n \geqslant 1$ 均有 $\tau_{(n)} \stackrel{\mathrm{d}}{=} \max\{\tau_{F_i}: 1 \leqslant i \leqslant n\} \leqslant \tau_G \leqslant \infty$.

由假定 3.1.1 知 $\tau_{(n)}$ 有极限(此时极限可以为无穷),不妨设 $\tau_0 = \lim_n \tau_{(n)}$, 显然 $\tau_0 \leqslant \tau_G \leqslant \infty$. 现在我们可以定义 $G(x)$ 的估计如下:

$$\begin{aligned} & g_n^{(2)}(x) \\ = & \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{h_n} K\left(\frac{x - x_i}{h_n}\right) \delta_i Z_i \hat{G}_{nS}^{-1}(Z_i) I[Z_i \leqslant M_n], \quad x \in [0,1]. \end{aligned} \tag{3.1.6}$$

其中 $M_n$ 为某小于 $\tau_{(n)}$ 且单调上升趋于 $\tau_0$ 的常数序列,而

$$\hat{G}_{nS}(t) = \begin{cases} \prod_{j=1}^{n} \left(\dfrac{N^+(Z_j)}{1 + N^+(Z_j)}\right)^{I[\delta_j = 0, Z_j \leqslant t]}, & \text{若 } t \leqslant Z_{(n)}, \\ 0, & \text{若 } t > Z_{(n)}, \end{cases}$$

此处 $Z_{(n)} = \max\{Z_1, Z_2, \cdots, Z_n\}$, $N^+(Z_j) = \sum_{i=1}^{n} I[Z_i \geqslant Z_j]$, $j = 1,2,\cdots,n$. $\hat{G}_{nS}$ 亦称为 $G_S$ 的 Kaplan-Meier 估计, 不过与通常 Kaplan-Meier 估计不同的是这里的删失变量(这里形式上把 $Y_1, Y_2, \cdots, Y_n$ 看做删失变量)独立非同分布.

本节 $C$ 可以表示任何与 $n$ 无关的正常数,还有 $N$ 可表示任何不同的正整数.

### 3.1.3 $g_n^{(1)}(\cdot)$ 的一些收敛性质

当 $G$ 已知时，我们用 $g_n^{(1)}(\cdot)$ 估计 $g(\cdot)$. 为研究 $g_n^{(1)}(\cdot)$ 的一些收敛性质，我们需要下面引理帮助.

**引理 3.1.1** 设 $\overline{K}(\cdot)$ 是 $\mathbf{R}^1$ 上连续函数，且存在 $M_0 \geqslant 0$，使 $u > M_0$ 时，$\overline{K}(u)$ 非增，$u < -M_0$ 时，$\overline{K}(u)$ 非减，$\int_{-\infty}^{\infty}\overline{K}(u)\,\mathrm{d}u < \infty$，若存在 $\Delta > 0$，使对所有 $n$ 均有 $\max\limits_{1\leqslant i\leqslant n}(x_i - x_{i-1}) \leqslant \Delta/n$，那么当 $n \to \infty$，$nh_n \to \infty$ 时，就有

$$h_n^{-1}\sum_{i=1}^{n}(x_i - x_{i-1})\overline{K}\left(\frac{x - x_i}{h_n}\right) \to \int_{-\infty}^{\infty}\overline{K}(u)\,\mathrm{d}u. \quad (3.1.7)$$

证明：由 $\overline{K}(u)$ 所满足的条件知必存在 $M_1 > M_0$，使得 $|u| > M_1$ 时，$\overline{K}(u) \geqslant 0$. 又由 $\int_{-\infty}^{\infty}\overline{K}(u)\,\mathrm{d}u < \infty$ 知 $\forall \varepsilon > 0$，$\exists M_2 > M_1 > M$ 使

$$\int_{M_2}^{\infty}\overline{K}(u)\,\mathrm{d}u < \frac{\varepsilon}{5}, \quad \int_{-\infty}^{-M_2}\overline{K}(u)\,\mathrm{d}u < \frac{\varepsilon}{5}.$$

对任取定的 $x \in (0,1)$（当 $x = 0$ 或 $x = 1$ 时可类似讨论），令 $u_i = \dfrac{x - x_i}{h_n}$，$i = 1,2,\cdots,n$，则由 $0 = x_0 < x_1 < \cdots < x_n = 1$ 得 $(x-1)h_n^{-1} = u_n < u_{n-1} < \cdots < u_0 = xh_n^{-1}$，再令 $v_i = u_{n-i}$，则得 $(x-1)h_n^{-1} = v_0 < v_1 < \cdots < v_n = xh_n^{-1}$，由于 $h_n \to 0$，故 $n$ 充分大时可保证 $(x-1)h_n^{-1} < -3M_2$，$xh_n^{-1} > 3M_2$，于是

$$h_n^{-1}\sum_{i=1}^{n}(x_i - x_{i-1})\overline{K}\left(\frac{x - x_i}{h_n}\right) = \sum_{i=1}^{n}(u_{i-1} - u_i)\overline{K}(u_i)$$

$$= \sum_{i=1}^{n}(v_i - v_{i-1})\overline{K}(v_{i-1})$$

$$= \sum_{\{i:(x-1)h_n^{-1}\leqslant v_i < -2M_2\}}(v_i - v_{i-1})\overline{K}(v_{i-1})$$

$$+ \sum_{\{i:-2M_2 \leqslant v_i \leqslant 2M_2\}} (v_i - v_{i-1})\overline{K}(v_{i-1})$$

$$+ \sum_{\{i:2M_2 < v_i \leqslant xh_n^{-1}\}} (v_i - v_{i-1})\overline{K}(v_{i-1})$$

$$\stackrel{\mathrm{d}}{=} Q_{1n} + Q_{2n} + Q_{3n}.$$

由 $|u| > M_2$ 时 $\overline{K}(u)$ 的单调性和非负性知

$$Q_{1n} \leqslant \int_{(x-1)h_n^{-1}}^{-2M_2} \overline{K}(u)\,\mathrm{d}u \leqslant \int_{-\infty}^{-M_2} \overline{K}(u)\,\mathrm{d}u < \frac{\varepsilon}{5},$$

$$Q_{3n} \leqslant \int_{M_2}^{xh_n^{-1}} \overline{K}(u)\,\mathrm{d}u \leqslant \int_{M_2}^{\infty} \overline{K}(u)\,\mathrm{d}u < \frac{\varepsilon}{5}.$$

若记 $i' = \min\{i: -2M_2 \leqslant v_i \leqslant 2M_2\}$, $i'' = \max\{i: -2M_2 \leqslant v_i \leqslant 2M_2\}$，则由 Rieman 积分的定义及 $\overline{K}(\cdot)$ 的有界性，易得

$$\left| Q_{2n} - \int_{-2M_2}^{2M_2} \overline{K}(u)\,\mathrm{d}u \right|$$

$$\leqslant \left| (v_{i'} + 2M_2)\overline{K}(-2M_2) + (2M_2 - v_{i''})\overline{K}(v_{i''}) + Q_{2n} - \int_{-2M_2}^{2M_2} \overline{K}(u)\,\mathrm{d}u \right|$$

$$+ (v_{i'} + 2M_2)\overline{K}(-2M_2) + (2M_2 - v_{i''})\overline{K}(v_{i''}) < \frac{\varepsilon}{5}.$$

综上即得

$$\left| h_n^{-1} \sum_{i=1}^{n} (x_i - x_{i-1})\overline{K}\left(\frac{x - x_i}{h_n}\right) - \int_{-\infty}^{\infty} \overline{K}(u)\,\mathrm{d}u \right|$$

$$\leqslant Q_{1n} + \left| Q_{2n} - \int_{-2M_2}^{2M_2} \overline{K}(u)\,\mathrm{d}u \right| + Q_{3n} + \int_{2M_2}^{\infty} \overline{K}(u)\,\mathrm{d}u$$

$$+ \int_{-\infty}^{-2M_2} \overline{K}(u)\,\mathrm{d}u < \varepsilon.$$

于是引理 3.1.1 得证.

为得到我们的结果，本节还假定 $\tau_{F_i} \leqslant \tau_G \leqslant \infty$， 且记 $a_{ni} =$

$\frac{x_i - x_{i-1}}{h_n}K\left(\frac{x - x_i}{h_n}\right)$, $Z_i^* = \delta_i Z_i G_S^{-1}(Z_i) - E\delta_i Z_i G_S^{-1}(Z_i)$, $Z_{in}^* = Z_i^* I[|a_{ni}Z_i^*| < n^{\frac{1}{p}}$, $\widetilde{K}(u) = |u|^{\alpha}K(u)$, $0 < \alpha < 1$, 有了这些准备,我们可得到下面的

**定理3.1.1** 设 $g(x)$ 在 $[0,1]$ 上满足 $\alpha$ 阶 Lipschitz 条件($0 < \alpha < 1$), $K(u)$ 为 $\mathbf{R}^1$ 上的连续概率密度核函数,且存在 $M_0 > 0$,当 $u > M_0$ 时 $\widetilde{K}(u)$ 非增,当 $u < -M_0$ 时 $\widetilde{K}(u)$ 非减,$\int_{-\infty}^{\infty}\widetilde{K}(u)\,du < \infty$,若存在 $\Delta > 0$,使对所有 $n$ 均有 $\max\limits_{1\leqslant i\leqslant n}(x_i - x_{i-1}) \leqslant \Delta/n$,且对某 $p > \sqrt{2}$,存在与 $n$ 无关的常数 $C$ 使 $\max\limits_{1\leqslant i\leqslant n} EY_i^p G_S^{1-p}(Y_i) \leqslant C$,则当

(i) $\sqrt{2} < p \leqslant 2$, $n^{\frac{2}{p}-p}h_n^{1-p} \to 0$,

(ii) $p > 2$, $n^{-\frac{p}{2}}h_n^{1-p} \to 0$

之一成立时,就有

$$g_n(x) \xrightarrow{P} g(x), \forall x \in [0,1].$$

注 3.1.1: $Y_i$ 服从均值为 $g(x_i)$,方差为 $\sigma_0^2 > 0$ 的截尾正态分布, $T_i$ 服从参数为 $\lambda$ 的指数分布时定理的条件 $\max\limits_{1\leqslant i\leqslant n} EY_i^p G_S{}^{1-p}(Y_i) \leqslant C$ 成立. 该条件成立的另一特例为 $\tau_{F_1} = \tau_{F_2} = \cdots = \tau_{F_n} = \tau_0 < \tau_G$,即 $Y_i$ 为具有相同上界的随机变量.

定理3.1.1的证明:由(3.1.3), $\forall x \in [0,1]$(注意 $x$ 一旦取定总认为不变),有

$$\begin{aligned} & g_n^{(1)}(x) - g(x) \\ = & \left[\sum_{i=1}^{n}\delta_i Z_i G_S^{-1}(Z_i)\frac{x_i - x_{i-1}}{h_n}K\left(\frac{x - x_i}{h_n}\right)\right. \\ & \left.- E\sum_{i=1}^{n}\delta_i Z_i G_S^{-1}(Z_i)\frac{x_i - x_{i-1}}{h_n}K\left(\frac{x - x_i}{h_n}\right)\right] \end{aligned}$$

$$+\left[\sum_{i=1}^{n} g(x_i)\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)-\int_{-\infty}^{\infty} g(x)K(u)\mathrm{d}u\right]$$

$$\stackrel{\mathrm{d}}{=} I_{1n}+I_{2n}. \tag{3.1.8}$$

我们先证 $I_{1n}\xrightarrow{P}0$，由于 $\forall\varepsilon>0$

$$P(|I_{1n}|>\varepsilon)=P\left(\left|\sum_{i=1}^{n} a_{ni}Z_i^*\right|>\varepsilon\right)$$

$$=P\left(\left|\sum_{i=1}^{n} a_{ni}Z_i^*\right|>\varepsilon,\text{存在 }1\leqslant i\leqslant n\text{ 使 }|a_{ni}Z_i^*|>n^{\frac{1}{p}}\right)$$

$$+P\left(\left|\sum_{i=1}^{n} a_{ni}Z_i^*\right|>\varepsilon,\text{对所有 }1\leqslant i\leqslant n\text{ 都有 }|a_{ni}Z_i^*|\leqslant n^{\frac{1}{p}}\right)$$

$$\stackrel{\mathrm{d}}{=} I_{11n}+I_{12n}. \tag{3.1.9}$$

由切比雪夫不等式

$$I_{11n}\leqslant\sum_{i=1}^{n}P(|a_{ni}Z_i^*|>n^{\frac{1}{p}})\leqslant n^{-1}\sum_{i=1}^{n}|a_{ni}|^pE|Z_i^*|^p, \tag{3.1.10}$$

又由 $C_r$ 不等式得

$$E|Z_i^*|^p\leqslant 2^{p-1}[E|\delta_iZ_iG_S^{-1}(Z_i)|^p+|E\delta_iZ_iG_S^{-1}(Z_i)|^p]$$

$$=2^{p-1}\left[\int_0^{\tau_{F_i}}\int_y^{\tau_G}|yG_S^{-1}(y)|^p\,\mathrm{d}G(t)\,\mathrm{d}F_i(y)+g^p(x_i)\right]$$

$$\leqslant 2^{p-1}[EY_i^pG_S^{1-p}(Y_i)+g^p(x_i)],\quad i=1,2,\cdots,n, \tag{3.1.11}$$

于是由定理有关条件即知

$$\max_{i\leqslant 1\leqslant n}E|Z_i^*|^p\leqslant C. \tag{3.1.12}$$

又

$$\sum_{i=1}^{n}a_{nk}^p\leqslant\max_{1\leqslant k\leqslant n}\left[\left(\frac{x_i-x_{i-1}}{h_n}\right)^{p-1}K^{p-1}\left(\frac{x-x_i}{h_n}\right)\right]$$

$$\cdot\sum_{k=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right), \tag{3.1.13}$$

由定理条件已知，$u>M_0$ 时 $\widetilde{K}(u)=|u|^{\alpha}K(u)$ 非增，从而 $u>$

$M_0$ 时 $K(u)$ 亦非增，又由 $u < -M_0$ 时 $\widetilde{K}(u) = |u|^\alpha K(u)$ 非减，知在 $u < -M_0$ 时 $K(u)$ 亦非减，再注意到 $K(u)$ 为连续概率密度核函数，故由引理 3.1.1 得

$$\sum_{k=1}^{n} \frac{x_i - x_{i-1}}{h_n} K\left(\frac{x - x_i}{h_n}\right) \to \int_{-\infty}^{\infty} K(u)\,\mathrm{d}u = 1 \qquad (3.1.14)$$

（注意本定理条件(i)，(ii)均蕴涵条件 $nh_n \to \infty$）. 由 $K(u)$ 所满足的条件容易看出 $K(u)$ 为 $\mathbf{R}^1$ 上的有界函数，由此及(3.1.14)，(3.1.13)和定理的有关条件知在 $p > \sqrt{2}$ 时

$$\sum_{i=1}^{n} a_{ni}^{p} \leqslant C(nh_n)^{1-p}, \qquad (3.1.15)$$

由(3.1.10)，(3.1.12)及(3.1.15)，得

$$0 \leqslant I_{11n} \leqslant C\, n^{-p} h_n{}^{1-p} \to 0, n \to \infty. \qquad (3.1.16)$$

而

$$\begin{aligned} I_{12n} &\leqslant P\left(\left|\sum_{i=1}^{n} a_{ni} Z_i^* I[|a_{ni} Z_i^*| \leqslant n^{\frac{1}{p}}]\right| > \varepsilon\right) \\ &\leqslant P\left(\left|\sum_{i=1}^{n} a_{ni} Z_i^* - \sum_{i=1}^{n} E a_{ni} Z_i^*\right| > \varepsilon - \left|\sum_{i=1}^{n} E a_{ni} Z_{in}^*\right|\right), \end{aligned} \qquad (3.1.17)$$

由于 $EZ_i^* = 0$，故

$$|E a_{ni} Z_{in}^*| = |E a_{ni} Z_i^* I[|a_{ni} Z_i^*| \geqslant n^{\frac{1}{p}}]|,$$

$i = 1,2,\cdots,n$，$Z_{in}^*$ 如本节一开始所定义. 于是对 $p > \sqrt{2}$，由(3.1.12)，(3.1.15)及条件(i)或(ii)得

$$\begin{aligned} \left|\sum_{i=1}^{n} E a_{ni} Z_{in}^*\right| &\leqslant \sum_{i=1}^{n} |E a_{ni} Z_i^* I[|a_{ni} Z_i^*| \geqslant n^{\frac{1}{p}}]| \\ &\leqslant \sum_{i=1}^{n} \{E[|a_{ni} Z_i^*| \cdot |a_{ni} Z_i^*|^{p-1} I[|a_{ni} Z_i^*| \\ &\geqslant n^{\frac{1}{p}}]) \Big/ n^{\frac{p-1}{p}}\} \end{aligned}$$

$$\leqslant n^{-1+\frac{1}{p}}\sum_{i=1}^{n}|a_{ni}|^{p}E|Z_i^*|^{p}\leqslant C\,n^{\frac{1}{p}-p}h_n^{1-p}\to 0,\quad(3.1.18)$$

由(3.1.18)知对上面的 $\varepsilon$,只要 $n$ 充分大就有

$$\Big|\sum_{i=1}^{n}Ea_{ni}Z_{in}^*\Big|<\varepsilon/2,\quad(3.1.19)$$

于是由(3.1.17),(3.1.19)知

$$I_{12n}\leqslant P\Big(\Big|\sum_{i=1}^{n}a_{ni}Z_{in}^*-\sum_{i=1}^{n}Ea_{ni}Z_{in}^*\Big|>\varepsilon/2\Big).\quad(3.1.20)$$

现在来证(3.1.20)的极限为零. 下面分$\sqrt{2}<p<2$和$p\geqslant 2$两种情况分别证明.

若$\sqrt{2}<p\leqslant 2$

$$\begin{aligned}&P\Big(\Big|\sum_{i=1}^{n}a_{ni}Z_{in}^*-\sum_{i=1}^{n}Ea_{ni}\Big|>\frac{\varepsilon}{2}\Big)\\&\leqslant CE\Big[\sum_{i=1}^{n}\Big(a_{ni}Z_{in}^*-\sum_{i=1}^{n}Ea_{ni}Z_{in}^*\Big)\Big]^2\\&\leqslant C\sum_{i=1}^{n}E|a_{ni}Z_{in}^*|^{2-p}\cdot|a_{ni}Z_{in}^*|^{p}\\&\leqslant C\,n^{\frac{2-p}{p}}\sum_{i=1}^{n}|a_{ni}|^{p}E|Z_i^*|^{p}\\&\leqslant Cn^{\frac{2}{p}-p}h_n^{1-p}\to 0,\end{aligned}\quad(3.1.21)$$

上面利用到(3.1.12),(3.1.15)与定理的条件(i).

若 $p\geqslant 2$,则由切比雪夫不等式及均值为零的独立随机变量和的不等式,$C_p$ 不等式及 Jensen 不等式得

$$\begin{aligned}&P\Big(\Big|\sum_{i=1}^{n}a_{ni}Z_i^*-\sum_{i=1}^{n}Ea_{ni}Z_i^*\Big|>\frac{\varepsilon}{2}\Big)\\&\leqslant CE\Big|\sum_{i=1}^{n}(a_{ni}Z_{in}^*-Ea_{ni}Z_{in}^*)\Big|^{p}\\&\leqslant Cn^{\frac{p}{2}-1}\sum_{i=1}^{n}E|a_{ni}Z_{in}^*-Ea_{ni}Z_{in}^*|^{p}\end{aligned}$$

$$\leqslant Cn^{\frac{p}{2}-1}\sum_{i=1}^{n}E\,|a_{ni}Z_{in}^{*}|^{p}$$

$$\leqslant C\,n^{\frac{p}{2}-1}\sum_{i=1}^{n}|a_{ni}|^{p}E\,|Z_{i}^{*}|^{p}\leqslant C\,n^{\frac{p}{2}-p}h_{n}^{\;1-p}\to 0. \quad (3.1.22)$$

由(3.1.21),(3.1.22)知在定理的条件(i)或条件(ii)下均有

$$P\left(\left|\sum_{i=1}^{n}a_{ni}Z_{i}^{*}-\sum_{i=1}^{n}Ea_{ni}Z_{i}^{*}\right|>\frac{\varepsilon}{2}\right)\to 0, \quad (3.1.23)$$

于是由(3.1.20),(3.1.23)知 $I_{12n}\to 0$,再由(3.1.16),(3.1.19)即得

$$I_{1n}\xrightarrow{p}0. \quad (3.1.24)$$

下面证 $I_{2n}\to 0(n\to\infty)$. 作变换

$$u_{i}=\frac{x-x_{i}}{h_{n}},i=1,2,\cdots,n, \quad (3.1.25)$$

则

$$\begin{aligned}I_{2n}&=\sum_{i=1}^{n}g(x_{i})\frac{x_{i}-x_{i-1}}{h_{n}}K\left(\frac{x-x_{i}}{h_{n}}\right)-g(x)\\&=-\sum_{i=1}^{n}(g(x-h_{n}u_{i})-g(x))(u_{i}-u_{i-1})K(u_{i})\\&\quad-\left[g(x)\sum_{i=1}^{n}(u_{i}-u_{i-1})K(u_{i})+g(x)\right]\\&\stackrel{\mathrm{d}}{=}-II_{1n}-II_{2n}.\end{aligned} \quad (3.1.26)$$

再应用变换(3.1.25),并应用引理3.1.1及(3.1.14),得

$$\begin{aligned}II_{2n}&=-g(x)\sum_{i=1}^{n}\frac{x_{i}-x_{i-1}}{h_{n}}K\left(\frac{x-x_{i}}{h_{n}}\right)+g(x)\\&\to -g(x)\int_{-\infty}^{\infty}K(u)\,\mathrm{d}u+g(x)\\&=0,\quad n\to\infty.\end{aligned} \quad (3.1.27)$$

又 $g(x)$ 满足 Lipschitz 条件,故

$$|II_{1}|\leqslant\left|\sum_{i=1}^{n}\right|g(x-h_{n}u_{i})-g(x)|(u_{i}-u_{i-1})K(u_{i})|$$

$$\leqslant C h_n^{\alpha} \left| \sum_{i=1}^{n} (u_i - u_{i-1}) \left| u_i \right|^{\alpha} K(u_i) \right|.$$

再一次应用上面所作的变换(3.1.25),得

$$\begin{aligned} |II_{1n}| &\leqslant Ch_n^{\alpha} \left| \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{h_n} \left| \frac{x - x_i}{h_n} \right|^{\alpha} K\left(\frac{x - x_i}{h_n}\right) \right| \\ &\leqslant Ch_n^{\alpha} \left| \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{h_n} \widetilde{K}\left(\frac{x - x_i}{h_n}\right) \right| \to 0, \quad n \to \infty. \end{aligned} \tag{3.1.28}$$

上面利用到 $\widetilde{K}$ 所满足的条件及引理 3.1.1. 于是(3.1.26),(3.1.27),(3.1.28)就证明了 $I_{2n} \to 0, n \to \infty$,由此及(3.1.8),(3.1.24)即知定理 3.1.1 的结论成立.

稍增大定理 3.1.1 中 $p$ 的下界,即可得下面的

**定理 3.1.2** 设 $g(\cdot)$, $K(\cdot)$, $\widetilde{K}(\cdot)$ 满足定理 3.1.1 同样的条件,且存在 $\Delta > 0$ 使得所有 $n$ 均有 $\max\limits_{1\leqslant i\leqslant n}(x_i - x_{i-1}) \leqslant \Delta/n$,若对某 $0 \leqslant \beta \leqslant 1$,当 $p > \dfrac{1-\beta+\sqrt{(1-\beta)^2+8}}{2}$ 时,有 $\max\limits_{1\leqslant i\leqslant n} EY_i^p G_S^{1-p}(Y_i) \leqslant C$,且 $\forall \varepsilon > 0$ 及任何固定的 $x \in [0,1]$,若

a) $0 < \beta \leqslant 1, \dfrac{1-\beta+\sqrt{(1-\beta)^2+8}}{2} < p \leqslant 2$,

$\sum\limits_{n=1}^{\infty} n^{\frac{2}{p}-p-\beta} h_n^{1-p} < \infty$,

b) $0 \leqslant \beta \leqslant 1, p > 2, \sum\limits_{i=1}^{\infty} n^{-\frac{p}{2}-\beta} h_n^{1-p} < \infty$

之一成立,则有

$$\sum_{n=1}^{\infty} n^{-\beta} P\left( \left| g_n^{(1)}(x) - g(x) \right| > \varepsilon \right) < \infty.$$

注 3.1.2:在定理 3.1.2 的条件 a) 中由于 $\dfrac{1-\beta+\sqrt{(1-\beta)^2+8}}{2} < p \leqslant 2$,因而 $\dfrac{2}{p} - p - \beta < -1$,此时只要取 $h_n$ 为充分慢地趋于零

的序列(如取 $h_n = \log^{-1} n$)就能保证

$$\sum_{n=1}^{\infty} n^{\frac{2}{p}-p-\beta} h_n^{1-\beta} < \infty.$$

定理 3.1.2 的证明：$\forall \varepsilon > 0$,及对任何取定的 $x \in [0,1]$

$$\begin{aligned}
&\sum_{n=1}^{\infty} n^{-\beta} P(|g_n^{(1)}(x) - g(x)| > \varepsilon) \\
&\leqslant \sum_{n=1}^{\infty} n^{-\beta} P\left(|g_n^{(1)}(x) - Eg_n^{(1)}(x)| > \frac{\varepsilon}{2}\right) \\
&\quad + \sum_{n=1}^{\infty} n^{-\beta} P\left(|Eg_n^{(1)}(x) - g(x)| > \frac{\varepsilon}{2}\right) \\
&\stackrel{\mathrm{d}}{=} E_{1n} + E_{2n}.
\end{aligned} \tag{3.1.29}$$

定理3.1.1已证 $|Eg_n^{(1)}(x) - g(x)| \to 0$,故对上述 $\varepsilon > 0$, $\exists N$ 使得 $n > N$ 时,有

$$|Eg_n^{(1)}(x) - g(x)| < \frac{\varepsilon}{2},$$

于是

$$E_{2n} = \sum_{n=1}^{N} n^{-\beta} P\left(|Eg_n^{(1)}(x) - g(x)| > \frac{\varepsilon}{2}\right) < \infty. \tag{3.1.30}$$

再按定理 3.1.1 证明 $I_{1n} \xrightarrow{p} 0$ 的方法,可得

$$E_{1n} \leqslant \sum_{n=1}^{\infty} n^{-\beta}(I_{11n} + I_{12n}), \tag{3.1.31}$$

其中 $I_{11n}, I_{12n}$ 如(3.1.9)所定义,从而由(3.1.16),(3.1.20),(3.1.21),(3.1.22)及(3.1.31)知必存在 $N \geqslant 1$,使

$$\begin{aligned}
E_{1n} &\leqslant \sum_{n=1}^{N} n^{-\beta}(I_{11n} + I_{12n}) \\
&\quad + C \sum_{n=N+1}^{\infty} n^{-\beta}\left[n^{-p} h_n^{1-p} + \max\left\{n^{\frac{2}{p}-p} h_n^{1-p}, n^{-\frac{p}{2}} h_n^{1-p}\right\}\right]
\end{aligned}$$

$$\leqslant\begin{cases}\sum_{n=1}^{N}\frac{(I_{11n}+I_{12n})}{n^{\beta}}+C\sum_{n=1}^{\infty}n^{\frac{2}{p}-p-\beta}h_n^{1-p}, \\ \quad 若\ 0<\beta\leqslant 1,\frac{1-\beta+\sqrt{(1-\beta)^2+8}}{2}<p\leqslant 2, \\ \sum_{n=1}^{N}\frac{(I_{11n}+I_{12n})}{n^{\beta}}+C\sum_{n=1}^{\infty}n^{-\frac{p}{2}-\beta}h_n^{1-p}, \\ \quad 若\ 0\leqslant\beta\leqslant 1,p>2.\end{cases} \tag{3.1.32}$$

于是由(3.1.29),(3.1.30),(3.1.32)及定理3.1.2的条件即知定理3.1.2的结果成立.

**推论 3.1.1** 设 $K(\cdot),g(\cdot),\widetilde{K}(\cdot)$ 均满足定理3.1.2的条件,且存在 $\Delta>0$ 使对所有 $n$ 均有 $\max\limits_{1\leqslant i\leqslant n}(x_i-x_{i-1})\leqslant\Delta/n$,若对某 $p>2$ 有 $\max\limits_{1\leqslant i\leqslant n}EY_i^pG_S^{1-p}(Y_i)\leqslant C$,则 $\forall\, x\in[0,1]$,当 $\sum\limits_{n=1}^{\infty}n^{-\frac{p}{2}}h_n^{1-p}<\infty$ 时,有

$$g_n^{(1)}(x)\xrightarrow{a.s.}g(x).$$

证明:在定理3.1.2中取 $\beta=0$,利用条件b)即可得

$$\sum_{n=1}^{\infty}P(|g_n^{(1)}(x)-g(x)|>\varepsilon)<\infty.$$

再利用Borel-Canteli引理即得本推论3.1.1的证明.

### 3.1.4 $g_n^{(2)}(x)$ 的收敛性质

本节除假定3.1.1成立外,还作如下假定:

假定3.1.2:$F_i,1\leqslant i\leqslant n,G$ 均为连续的分布函数.

假定3.1.3:$\tau_0<\tau_G$($\tau_0$ 如节3.1.2所定义).

假定3.1.4:对某 $b>0,\liminf\limits_{n}\log n\bar{F}_S(M_n)\geqslant b$.

在给出结果之前,我们先证

**引理 3.1.2** 在假定3.1.1-3.1.4下,$\forall\,\varepsilon>0$,有

$$\sum_{n=1}^{\infty} P\left(\sup_{0\leqslant t\leqslant M_n} \left|\log\hat{G}_{nS}(t) - \log G_S(t)\right| > \varepsilon\right) < \infty. \tag{3.1.33}$$

证明：$\forall\, t \in [0, M_n]$，定义 $\beta_j(t) = I[\delta_j = 0, Z_j \leqslant t]$，则

$$\log\hat{G}_{nS}(t) = \sum_{j=1}^{n} I[\delta_j = 0, Z_j \leqslant t]\log\{N^+(Z_j)/(N^+(Z_j)+1)\}.$$

对 $\log\left(1 - \dfrac{1}{N^+(Z_j)+1}\right)$进行 Taylor 展开，即可得

$$\begin{aligned}
&\log\hat{G}_{nS}(t) - \log G_S(t) \\
&= \left[-\frac{1}{n}\sum_{j=1}^{n}\beta_j(t)\overline{H}_S^{-1}(Z_j) - \log G_S(t)\right] \\
&\quad + \left[-\sum_{j=1}^{n}\beta_j(t)\sum_{l=2}^{n}\frac{1}{l}(1+N^+(Z_j))^{-l}\right] \\
&\quad + \left[-\frac{1}{n}\beta_j(t)\{n(1+N^+(Z_j))^{-1} - \overline{H}_S^{-1}(Z_j)\}\right] \\
&\stackrel{\mathrm{d}}{=} R_{n1} + R_{n2} + R_{n3}.
\end{aligned} \tag{3.1.34}$$

因而 $\forall\, \varepsilon > 0$，对一切 $n \geqslant 1$ 及任何 $t \in [0, M_n]$，有

$$\begin{aligned}
&P(|\log\hat{G}_{nS}(t) - \log G_S(t)| > \varepsilon) \\
&\leqslant \sum_{i=1}^{3} P\left(|R_{ni}(t)| > \frac{\varepsilon}{3}\right).
\end{aligned} \tag{3.1.35}$$

由于 $-E\beta_j(t)\overline{H}_S^{-1}(Z_j) = -\int_0^t \dfrac{F_{jS}(u)}{\overline{H}_S(u)}\,\mathrm{d}G(u)$，于是

$$\frac{1}{n}\sum_{j=1}^{n} E[-\beta_j(t)\overline{H}_S(Z_j)] = \log G_S(t),$$

因而

$$R_{n1}(t) = \frac{1}{n}\sum_{j=1}^{n}[(-\beta_j(t)\overline{H}_S(Z_j)) - E(-\beta_j(t)\overline{H}_S(Z_j))]. \tag{3.1.36}$$

利用切比雪夫不等式，均值为零的独立随机变量和的绝对矩不等式，$C_p$ 不等式，Jenssen 不等式知 $p > 2$ 时，对一切 $n \geqslant 1$ 及

$\forall t \in [0, M_n]$, 有

$$P\left(|R_{n1}(t)| > \frac{\varepsilon}{3}\right)$$

$$\leqslant Cn^{-p}E\left|\sum_{j=1}^{n}(\beta_j(t)\overline{H}_S^{-1}(Z_j)) - E\beta_j(t)\overline{H}_S^{-1}(Z_j)]\right|^p$$

$$\leqslant Cn^{-p}\cdot n^{\frac{p}{2}-1}\sum_{j=1}^{n}E\left|\beta_j(t)\overline{H}_S^{-1}(Z_j) - E\beta_j(t)\overline{H}_S^{-1}(Z_j)\right|^p$$

$$\leqslant Cn^{-\frac{p}{2}-1}\sum_{j=1}^{n}E\left|\beta_j(t)\overline{H}_S^{-1}(Z_j)\right|^p$$

$$\leqslant C\,G_S^{-p}(\tau_0)n^{-\frac{p}{2}}\overline{F}_S^{1-p}(M_n) \leqslant C\,n^{-\frac{p}{2}}\overline{F}_S^{1-2p}(M_n). \quad (3.1.37)$$

由假定 3.1.4 及(3.1.37) 知存在 $N$, 使得当 $n > N$ 时, 对一切 $t \in [0, M_n]$, 有

$$P\left(|R_{n1}(t)| > \frac{\varepsilon}{3}\right) \leqslant Cb^{2p-1}n^{\frac{p}{2}}\log^{2p-1}n \leqslant Cn^{-\frac{p}{2}}\log^{2p-1}n. \quad (3.1.38)$$

又

$$\sum_{l=2}^{\infty}\frac{1}{l}(1 + N^+(Z_j))^{-l} \leqslant \frac{1}{(1 + N^+(Z_j))^2}, \quad (3.1.39)$$

故对一切 $n$ 及 $\forall t \in [0, M_n]$,有

$$P\left(|R_{n2}(t)| > \frac{\varepsilon}{3}\right)$$

$$\leqslant P\left(\left|\sum_{j=1}^{n}\frac{\beta_j(t)}{(1 + N^+(Z_j))^2}\right| > \frac{\varepsilon}{3}\right)$$

$$\leqslant CE\left|\sum_{j=1}^{n}\frac{\beta_j(t)}{(1 + N^+(Z_j))^2}\right|^p$$

$$\leqslant Cn^{p-1}\sum_{j=1}^{n}E\{\beta_j(t)E[(1 + N^+(Z_j))^{-2p}|(Z_j, \delta_j)]\}. \quad (3.1.40)$$

由 Koul, Susarlar 与 Ryzin[47] 的推论 7.1 知

$$E[(1+N^+(Z_j))^{-2p}|(Z_j,\delta_j)]\leqslant n^{-2p}[\overline{H}_S(Z_j)-n^{-1}]^{-2p}. \tag{3.1.41}$$

而由假定3.1.3与3.1.4知当 $n$ 大于某 $N$ 时,对一切 $t\in[0,M_n]$ 均有 $\overline{H}_S(t)-n^{-1}\geqslant\overline{H}_S(M_n)-n^{-1}>0$,且

$$\begin{aligned}(\overline{H}_S(t)-n^{-1})^{-1}&\leqslant\frac{[1+(n\overline{H}_S(M_n)-1)^{-1}]}{\overline{H}_S(t)}\\&\leqslant[1+(a-1)^{-1}]\overline{H}_S^{-1}(t),\end{aligned} \tag{3.1.42}$$

其中 $a$ 为大于1的某常数. 于是由(3.1.40),(3.1.41)与(3.1.42)得 $n$ 大于某 $N$ 时,对一切 $t\in[0,M_n]$

$$\begin{aligned}&P\left(|R_{n2}(t)|>\frac{\varepsilon}{3}\right)\\&\leqslant Cn^{-p-1}\sum_{j=1}^{n}\int_0^t\frac{F_{jS}(u)}{(\overline{H}_S(u)-n^{-1})^{2p}}\,\mathrm{d}G(u)\\&\leqslant Cn^{-p}G_S^{-2p}(\tau_0)\overline{F}_S^{1-2p}(M_n)\leqslant Cn^{-p}\log^{2p-1}n.\end{aligned} \tag{3.1.43}$$

又由切比雪夫不等式及 $C_p$ 不等式知对任何 $n$ 及 $\forall t\in[0,M_n]$,有

$$\begin{aligned}&P\left(|R_{n3}(t)|>\frac{\varepsilon}{3}\right)\\&\leqslant Cn^{-1}\sum_{j=1}^{n}E\left\{\beta_j(t)E\left[\left|\frac{n}{1+N^+(Z_j)}-\frac{1}{\overline{H}_S(Z_j)}\right|^p|(Z_j,\delta_j)\right]\right\}.\end{aligned} \tag{3.1.44}$$

利用 Cauchy-Schwartz 不等式,得

$$\begin{aligned}&E\left[\left|\frac{n}{1+N^+(Z_j)}-\frac{1}{\overline{H}_S(Z_j)}\right|^p|(Z_j,\delta_j)\right]\\&\leqslant\overline{H}^{-p}(Z_j)E^{\frac{1}{2}}[(1+N^+(Z_j))^{-2p}|(Z_j,\delta_j)]E^{\frac{1}{2}}\\&\quad\cdot[|(1+N^+(Z_j))-n\overline{H}_S(Z_j)|^{2p}|(Z_j,\delta_j)]\end{aligned} \tag{3.1.45}$$

利用中心化独立贝努里随机变量的绝对矩不等式(即均值为零的独立随机变量和的不等式可得

$$E\left[\left|(1+N^{+}(Z_j))-n\overline{H}_S(Z_j)\right|^{2p}\left|(Z_j,\delta_j)\right.\right]$$

$$\leqslant Cn^{p-1}\sum_{j=1}^{n}\left[H_{jS}(Z_j)+H_{jS}^{2p}(Z_j)\right]\leqslant Cn^{p}, \qquad (3.1.46)$$

于是由(3.1.44),(3.1.45),(3.1.41)与(3.1.46)并注意应用(3.1.43)的方法,得 $n$ 大于某$N$,对一切 $t\in[0,M_n]$,都有

$$P\left(\left|R_{n3}(t)>\frac{\varepsilon}{3}\right.\right)$$

$$\leqslant Cn^{-\frac{p}{2}-1}\sum_{j=1}^{n}E\beta_j(t)\left[\overline{H}_S^{-p}(Z_j)(\overline{H}_S(Z_j)-n^{-1})^{-p}\right]$$

$$\leqslant Cn^{-\frac{p}{2}}G_S^{-p}(\tau_0)\int_0^{M_n}\overline{F}_S^{1-2p}(u)\,\mathrm{d}G(u)$$

$$\leqslant Cn^{-\frac{p}{2}}\log^{2p-1}n. \qquad (3.1.47)$$

综合(3.1.35),(3.1.38),(3.1.43)与(3.1.47)即知对某 $N$,当 $n>N$ 时

$$P(\left|\log\hat{G}_{nS}(t)-\log G_S(t)\right|>\varepsilon)\leqslant Cn^{-\frac{p}{2}}\log^{2p-1}n \qquad (3.1.48)$$

对一切 $t\in[0,M_n]$ 成立(注意上面的 $p>2$).

又由假定3.1.2,3.1.3知对任何 $n$,$\log G_S(t)$ 在$[0,\tau_0]$上连续,因而一致连续. 由于$[0,M_n]\subset[0,\tau_0]$,由此对上述 $\varepsilon>0$,$\exists\,\delta>0$ 使对一切 $n\geqslant 1$,$[0,M_n]$ 中任意两点 $\eta_i,\eta_j$,只要 $|\eta_i-\eta_j|<\delta$,就有 $|\log G_S(\eta_i)-\log G_S(\eta_j)|\leqslant\frac{\varepsilon}{2}$. 现在在$[0,M_n]$中插入$\left[\frac{M_n}{\delta}\right]$个等分点 $\eta_1<\eta_2<\cdots<\eta_{N(\delta)}$,此处 $N(\delta)\stackrel{\mathrm{d}}{=}\left[\frac{M_n}{\delta}\right]$. 再记

$\eta_0=0,\eta_{N(\delta)+1}=M_n$,由于 $|\eta_i-\eta_{i-1}|\leqslant M_n/\left(\left[\frac{M_n}{\delta}\right]+1\right)<\delta$,故 $|\log G_S(\eta_i)-\log G_S(\eta_{i-1})|\leqslant\varepsilon/2, i=1,2,\cdots,N(\delta)+1$. 若 $|\log\hat{G}_{nS}(\eta_i)-\log G_S(\eta_i)|\leqslant\varepsilon/2$, 且 $|\log\hat{G}_{nS}(\eta_{i-1})-\log G_S(\eta_{i-1})|\leqslant\varepsilon/2$,则对 $\forall\, t\in[\eta_{i-1},\eta_i]$ 必有 $|\log\hat{G}_{nS}(t)-$

$\log G_S(t)| \leqslant \varepsilon$，由此知若 $\sup_{0\leqslant t\leqslant M_n} |\log\hat{G}_{nS}(t) - \log G_S(t)| > \varepsilon$，则必存在 $i$，使得 $|\log\hat{G}_{nS}(\eta_i) - \log G_S(\eta_i)| > \varepsilon$，由此及(3.1.48)知 $n > N$ 时

$$
\begin{aligned}
&P(\sup_{0\leqslant t\leqslant M_n} |\log\hat{G}_{nS}(t) - \log G_S(t)| > \varepsilon) \\
\leqslant\ & P(\bigcup_{i=0}^{N(\delta)+1} \{|\log\hat{G}_{nS}(\eta_i) - \log G_S(\eta_i)| > \varepsilon\}) \\
\leqslant\ & (N(\delta) + 2) \max_{0\leqslant i\leqslant N(\delta)+1} P(|\log\hat{G}_{nS}(\eta_i) - \log G_S(\eta_i)| > \varepsilon) \\
\leqslant\ & C\left(\frac{\tau_0}{\delta} + 2\right) n^{-\frac{p}{2}}\log^{-\frac{p}{2}} n \leqslant C\, n^{-\frac{p}{2}}\log^{2p-1} n.
\end{aligned}
\tag{3.1.49}
$$

因而当 $p > 2$ 时就有

$$
\begin{aligned}
&\sum_{n=1}^{\infty} P(\sup_{0\leqslant t\leqslant M_n} |\log\hat{G}_{nS}(t) - \log G_S(t)| > \varepsilon) \\
=\ & \sum_{n=1}^{N} P(\sup_{0\leqslant t\leqslant M_n} |\log\hat{G}_{nS}(t) - \log G_S(t)| > \varepsilon) \\
&+ \sum_{n=N+1}^{\infty} P(\sup_{0\leqslant t\leqslant M_n} |\log\hat{G}_{nS}(t) - \log G_S(t)| > \varepsilon) < \infty.
\end{aligned}
$$

这就证明了引理 3.1.2.

**定理 3.1.3** 设 $g(\cdot), K(\cdot), \widetilde{K}(\cdot)$ 满足定理 3.1.1 的条件，且引理 3.1.2 的假设全部成立，若存在 $\Delta > 0$ 使对所有 $n$ 均有 $\max_{1\leqslant i\leqslant n}(x_i - x_{i-1}) \leqslant \Delta/n$，则 $\forall\, \varepsilon > 0$ 及 $\forall\, x \in [0,1]$，当 $F_i, i = 1, 2, \cdots$ 在 $\tau_0$ 等度连续，且对某 $p > 2$，$\sum_{n=1}^{\infty} n^{-\frac{p}{2}} h_n^{1-p} < \infty$ 时，有

$$
\sum_{n=1}^{\infty} P(|g_n^{(2)}(x) - g(x)| > \varepsilon) < \infty.
$$

证明：对任意取定的 $x \in [0,1]$.

$$
\begin{aligned}
&|g_n^{(2)}(x) - g(x)| \\
\leqslant\ & \left| \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{h_n} K\left(\frac{x - x_i}{h_n}\right) Z_i\delta_i\hat{G}_{nS}^{-1}(Z_i) I[Z_i \leqslant M_n]\right.
\end{aligned}
$$

$$
\begin{aligned}
&-\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)Z_i\delta_iG_S^{-1}(Z_i)I[Z_i\leqslant M_n]\Bigg|\\
&+\Bigg|\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)Z_i\delta_iG_S^{-1}(Z_i)\\
&-\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)E[Z_i\delta_iG_S^{-1}(Z_i)]\Bigg|\\
&+\Bigg|\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)E[Z_i\delta_iG_S^{-1}(Z_i)]-g(x)\Bigg|\\
&+\Bigg|\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)Z_i\delta_iG_S^{-1}(Z_i)I[Z_i>M_n]\\
&-\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)E(Z_i\delta_iG_S^{-1}(Z_i)I[Z_i>M_n])\Bigg|\\
&+\Bigg|\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)E(Z_i\delta_iG_S^{-1}(Z_i)I[Z_i>M_n])\Bigg|\\
&\stackrel{\mathrm{d}}{=}\sum_{i=1}^{5}\Delta_{in}.
\end{aligned}
\tag{3.1.50}
$$

由于 $\tau_0<\tau_G$,因而 $\max\limits_{1\leqslant i\leqslant n}EY_i^{p}G_S^{1-p}(Y_i)\leqslant\tau_0^{P}G_S^{1-p}(\tau_0)\leqslant C$,于是我们可利用节 3.1.3 中证明 $I_{1n}\xrightarrow{p}0$ 时的一些方法得在 $p>2$, $n$ 充分大时,有

$$
P\left(\Delta_{2n}>\frac{\varepsilon}{5}\right)\leqslant Cn^{-\frac{p}{2}}h_n^{1-p},\tag{3.1.51}
$$

$$
P\left(\Delta_{4n}>\frac{\varepsilon}{5}\right)\leqslant Cn^{-\frac{p}{2}}h_n^{1-p}.\tag{3.1.52}
$$

又由(3.1.14)知 $n$ 充分大时

$$
\begin{aligned}
\Delta_{5n}&\leqslant(\max_{1\leqslant i\leqslant n}E\delta_iZ_iG_S^{-1}(Z_i)I[Z_i>M_n])\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{h_n}K\left(\frac{x-x_i}{h_n}\right)\\
&\leqslant C\max_{1\leqslant i\leqslant n}\int_{M_n}^{\tau_{F_i}}y\,\mathrm{d}F_i(y)\\
&\leqslant C\max_{1\leqslant i\leqslant n}[F_i(\tau_0)-F_i(M_n)]\to0.
\end{aligned}
\tag{3.1.53}
$$

上面(3.1.53)中最后的极限式用到 $F_i, i=1,2,\cdots$ 在 $\tau_0$ 的等度连续性(即 $F_i(\cdot)$ 在 $\tau_0$ 连续对所有 $i \geqslant 1$ 一致). 因而对上述 $\varepsilon$, 当 $n$ 充分大时

$$P\left(\Delta_{5n} > \frac{\varepsilon}{5}\right) = 0. \tag{3.1.54}$$

由(3.1.8)所定义的 $I_{2n} \to 0$ 的证明,知在此亦有 $\Delta_{3n} \to 0$,因而对上述 $\varepsilon$, 当 $n$ 充分大时

$$P\left(\Delta_{3n} > \frac{\varepsilon}{5}\right) = 0. \tag{3.1.55}$$

而由(3.1.14)及引理3.1.2并注意到 $M_n < \tau_0 < \tau_G$,知 $n$ 充分大时

$$\begin{aligned}
&P\left(\Delta_{1n} > \frac{\varepsilon}{5}\right) \\
&\leqslant P\left(\sum_{i=1}^{n} \frac{x_i - x_{i-1}}{h_n} K\left(\frac{x - x_i}{h_n}\right) \mid Z_i \delta_i (\hat{G}_{nS}^{-1}(Z_i) \right. \\
&\quad \left. - G_S^{-1}(Z_i)) I[Z_i \leqslant M_n] \mid > \frac{\varepsilon}{5}\right) \\
&\leqslant P\left(\tau_0 \sup_{0 \leqslant t \leqslant M_n} |\hat{G}_{nS}^{-1}(t) - G_S^{-1}(t)| \sum_{i=1}^{n} \frac{x_i - x_{i-1}}{h_n} K\left(\frac{x - x_i}{h_n}\right) > \frac{\varepsilon}{5}\right) \\
&\leqslant P\left(\sup_{0 \leqslant t \leqslant M_n} |e^{\log \hat{G}_{nS}(t)} - e^{\log G_S^{-1}(t)}| > \frac{\varepsilon}{10\tau_0}\right) \\
&\leqslant P(G_S^{-1}(\tau_0) \sup_{0 \leqslant t \leqslant M_n} |e^{\theta(\log \hat{G}_{nS}^{-1}(t) - \log G_S^{-1}(t))} (\log \hat{G}_{nS}^{-1}(t) \\
&\quad - \log G_S^{-1}(t))| > \frac{\varepsilon}{10\tau_0}) \\
&\leqslant P(e^{\sup_{0 \leqslant t \leqslant M_n} |\log \hat{G}_{nS}(t) - \log G_S(t)|} \sup_{0 \leqslant t \leqslant M_n} |\log \hat{G}_{nS}(t) \\
&\quad - \log G_S(t)| > \frac{G_S(t_0)\varepsilon}{10\tau_0}) \\
&= P\left(e^{\Delta_n} \Delta_n > \frac{\varepsilon G_S(\tau_0)}{10\tau_0}, \Delta_n > 1\right) + P\left(e^{\Delta_n} \Delta_n > \frac{\varepsilon G_S(\tau_0)}{10\tau_0}, \Delta_n \leqslant 1\right)
\end{aligned}$$

$$\leqslant P(\Delta_n > 1) + P\left(\Delta_n > \frac{\varepsilon G_S(\tau_0)}{10\tau_0 \mathrm{e}}\right). \qquad (3.1.56)$$

上式中 $0 < \theta < 1, \Delta_n \stackrel{\mathrm{d}}{=} \sup\limits_{0 \leqslant t \leqslant M_n} |\log \hat{G}_{nS}(t) - \log G_S(t)|$. 由于

$$P(|g_n^{(2)}(x) - g(x)| > \varepsilon) \leqslant \sum_{i=1}^{5} P\left(\Delta_{in} > \frac{\varepsilon}{5}\right), \qquad (3.1.57)$$

于是由(3.1.51),(3.1.52),(3.1.54),(3.1.55),(3.1.56) 及引理 3.1.2 与定理有关条件即得本定理的结论.

**推论 3.1.2** 若定理 3.1.3 的条件全部满足,则对 $\forall x \in [0, 1]$,有

$$g_n^{(2)}(x) \xrightarrow{a.s.} g(x).$$

证明:利用定理 3.1.3 的结果及 Borel-Cantelli 引理即得本推论.

## §3.2 半参数回归模型中的相合估计

### 3.2.1 记号与主要结果

考虑固定设计下半参数回归模型

$$Y_i = x_i\beta + g(t_i) + \varepsilon_i, \quad i = 1,2,\cdots,n, \qquad (3.2.1)$$

其中 $x_i, t_i (i = 1,2,\cdots,n)$ 是已知的设计点列,$\beta$ 是一维未知参数,$g(\cdot)$ 是定义在闭区间 $I$ 上的未知回归函数,$\varepsilon_j, j = 1,2,\cdots,n$,是均值为零方差为 $\sigma^2$ 的随机误差序列.

关于模型(3.2.1),有几种估计 $\beta$ 与 $g(\cdot)$ 的方法. 其中一个主要的方法是 Engle,Granger 与 Rice[29] 所提出的偏最小二乘法,胡舒合[102] 在随机误差为 $\phi$ 混合情形下定义了 $\beta$ 与 $g(\cdot)$ 的估计并研究了其强相合性. 而对 $x_i, t_i, i = 1,2,\cdots,n$ 为随机的场合已有更多的研究(见洪圣岩[101],Chen[20],高集体与赵林城[98],高集体[99] 及其所列文献).

众所周知,在一些实际问题,如可靠性寿命试验,医药追踪试验及对生存分析等领域的研究中,$Y_i$ 常常因被某随机变量 $C_i$ 随机右截断而不能被完全观察,我们仅能观察到

$$Z_i = \min(Y_i, C_i), \quad \delta_i = I[Y_i \leqslant C_i],$$

其中 $i = 1,2,\cdots,n$, $I[\cdot]$ 表示某事件的示性函数.

现在的问题是要考虑如何利用上面所获的截断数据 $(Z_i, \delta_i)$, $i = 1,2,\cdots,n$ 去估计参数 $\beta$ 和未知回归函数 $g(\cdot)$,并希望所定义的估计有通常的大样本性质.

以下,我们均假设 $Y_1, Y_2, \cdots, Y_n$ 彼此独立,且 $Y_i$ 有连续的分布函数 $F_i$, $i = 1,2,\cdots,n$. $C_1, C_2, \cdots, C_n$ 独立同分布,有共同的连续分布函数 $G$. 易见 $Z_i$ 的分布函数是 $H_i = (1 - F_i)(1 - G)$, $i = 1,2,\cdots,n$. 此外,为方便计,对任何分布函数 $V(\cdot)$,定义 $V_S(\cdot) = 1 - V(\cdot)$, $\tau_V = \inf\{t : V(t) = 1\}$,对任何 $r > 0$, $V^{-r}(\cdot) = [V(\cdot)]^{-r}$,且约定本节中的 $\alpha$ 可表示任何所需的常数.

以下始终假定 $\tau_{F_i} \leqslant \tau_G$, $i = 1,2,\cdots,n$. $g(t)$ 在 $I$ 上连续.

注意到

$$E\delta_i Z_i G_S{}^{-1}(Z_i) = \int_{-\infty}^{\tau_{F_i}} \int_y^{\tau_G} \frac{y}{1 - G(y)} \mathrm{d}G(t) \mathrm{d}F_i(y) = EY_i$$
$$= \beta x_i + g(t_i), \quad i = 1,2,\cdots,n. \tag{3.2.2}$$

由此我们认为 $\{\delta_i Z_i G_S^{-1}(Z_i)\}$ 遵从如下模型:

$$\delta_i Z_i G_S^{-1}(Z_i) = \beta x_i + g(t_i) + e_i, \; i = 1,2,\cdots,n, \tag{3.2.3}$$

其中 $e_1, e_2, \cdots, e_n$ 独立且均值为零.

记 $Z_{iG} = \delta_i Z_i G_S^{-1}(Z_i)$, $i = 1,2,\cdots,n$. 于是 $G$ 已知时,仿胡舒合[102]中的方法在此构造 $g(\cdot)$ 的估计如下

$$g_n(t) \overset{\mathrm{d}}{=} g_n(t, \beta) = \sum_{j=1}^{n} W_{nj}(t)(Z_{jG} - x_j \beta). \tag{3.2.4}$$

其中 $0 \leqslant W_{nj}(t) \leqslant 1$ 为权函数, $j = 1,2,\cdots,n$.

又记 $\tilde{Z}_{iG} = Z_{iG} - \sum_{j=1}^{n} W_{nj}(t_i) Z_{jG}$, $\tilde{x}_i = x_i - \sum_{j=1}^{n} W_{nj}(t_i) x_j$, $\tilde{S}_n{}^2$

$=\sum_{i=1}^{n}\tilde{x}_i^2$. 于是又参考非截尾情形下对 $\beta$ 估计的定义,在此我们可定义 $\beta$ 的估计

$$\tilde{\beta}_n=\sum_{i=1}^{n}\tilde{x}_i\tilde{Z}_{iG}/\tilde{S}_n^2. \tag{3.2.5}$$

因而由(3.2.4),(3.2.5)知 $g(\cdot)$ 在 $G$ 已知的情形的最终估计为

$$\tilde{g}_n(t)\stackrel{\mathrm{d}}{=}g_n(t,\tilde{\beta}_n)=\sum_{j=1}^{n}W_{nj}(t)(Z_{jG}-x_j\tilde{\beta}_n). \tag{3.2.6}$$

首先,我们给出 $\tilde{\beta}_n,\tilde{g}_n(\cdot)$ 的强相合性和 $p(\geqslant 2)$ 阶平均相合定理,为叙述简洁,先列举如下条件

A1:存在 $p\geqslant 2$,使得 $E|Y_i|^pG_S^{1-p}(Y_i)\leqslant\alpha, i=1,2,\cdots,n$.

A2:(i) $\tilde{S}_n^{-2}\sum_{i=1}^{n}|\tilde{x}_i|\leqslant\alpha$.

(ii) $\tilde{S}_n^{-1}\leqslant\alpha n^{-\frac{1}{2}}\sqrt{\log n}$.

(iii) 对 A1 中的 $p\geqslant 2$, $\tilde{S}_n^{-1}\leqslant\alpha\min\{n^{-1/2}(\log n)^{-1}, n^{-2/p}(\log n)^{-1}\}$.

A3:(i) 对 $t\in I$ 一致地有 $\left|\sum_{j=1}^{n}W_{nj}(t)-1\right|=o(1)$ 成立.

(ii) 对任何 $a>0$, $\sum_{\{j:|t_j-t|>a\}}W_{nj}(t)=o(1)$ 关于 $t\in I$ 一致成立.

(iii) 对 $t\in I$ 一致地有

$$\max_{1\leqslant j\leqslant n}W_{nj}(t)\leqslant\alpha\min\{n^{-1/2}(\log n)^{-1}, n^{-2/p}(\log n)^{-1}\}$$

成立.

(iv) $\left|\sum_{j=1}^{n}W_{nj}(t)x_j\right|\leqslant\alpha_t$,对 $t\in I$ 成立,这里 $\alpha_t$ 表示可与 $t$ 有关的常数(就取定 $t$ 而言).

**定理 3.2.1** 若条件 A1,A2(i),(iii)与 A3(i),(ii),(iii)满足,则

$$\tilde{\beta}_n \xrightarrow{a.s.} \beta. \tag{3.2.7}$$

若再满足 A3(iv),则

$$\tilde{g}_n(t) \xrightarrow{a.s.} g(t), \forall t \in I. \tag{3.2.8}$$

**定理 3.2.2** 若条件 A1,A2(ii) 与 A3(i),(ii),(iii) 满足,则

$$E|\tilde{\beta}_n - \beta|^p \to 0, n \to \infty. \tag{3.2.9}$$

若再满足 A3(iv),则

$$E|\tilde{g}_n(t) - g(t)|^p \to 0, n \to \infty. \tag{3.2.10}$$

但在 $G$ 未知时,为估计 $\beta$ 与 $g(\cdot)$,一个自然的作法是用 $G$ 的估计取代 $\tilde{\beta}$, $\tilde{g}_n(\cdot)$ 中的 $G$. 这里我们采用 Koul, Susarla 与 Ryzin[47] 中

$$\tilde{G}_{ns}(t) = \prod_{j=1}^{n} \{(1 + Z^+(Z_j))/(2 + N^+(Z_j))\}^{I[\delta_j = 0, Z_j \leqslant t]},$$
$$-\infty < t < \infty$$

作为 $G_S$ 的估计,其中

$$N^+ = \sum_{i=1}^{n} I[Z_i > \cdot].$$

但考虑到 $\hat{G}_{ns}(t)$ 的渐近方差在充分大 $t$ 时的"爆炸"行为,因而这里我们并不简单地用 $\hat{G}_{nS}$ 取代 $\tilde{\beta}$, $\tilde{g}$ 中的 $G_S$ 以获 $\beta$, $g$ 的估计,而是作适当的修正,即定义

$$\hat{\beta}_n = \sum_{i=1}^{n} \tilde{x}_i \{ Z_{i\hat{G}_n} I[Z_i \leqslant M_n]$$
$$- \sum_{j=1}^{n} W_{nj}(t_i) Z_{j\hat{G}_n} I[Z_j \leqslant M_n] \} / \tilde{S}_n^2, \tag{3.2.11}$$

$$\hat{g}_n(t) = \sum_{j=1}^{n} W_{nj}(t) \{ Z_{j\hat{G}_n} I[Z_j \leqslant M_n] - x_j \hat{\beta} \}, \forall t \in I \tag{3.2.12}$$

分别作为 $\beta$ 与 $g$ 的估计,其中 $Z_{i\hat{G}_n}$ 是用 $\hat{G}_{nS}$ 取代 $Z_{iG}$ 中的 $G_S$ 而得到,

$M_n$ 是趋于 $\tau_0 \stackrel{d}{=} \lim_n \tau_{(n)}$，的正常数序列，此处 $\tau_{(n)} \stackrel{d}{=} \max\{\tau_{F_i}:1 \leqslant i \leqslant n\}$. 为给出 $\hat{\beta}_n, \hat{g}_n$ 的强相合及 $p(\geqslant 2)$ 阶平均相合定理，我们再列下面条件：

A4：(i) 对 $b > 0$ 及 A1(i) 中的 $p \geqslant 2, \lim_n \log n\overline{H}_S^{3p}(M_n) \geqslant b$，

其中 $\overline{H}_S(\cdot) = \frac{1}{n}\sum_{i=1}^{n} H_{iS}(\cdot)$.

(ii) 对 $0 < \theta < 1$ 及 A1(i) 中 $p \geqslant 2$，存在 $d > 0$，使得 $\lim_n n^{\theta}\overline{H}_S^{3p}(M_n) \geqslant d$ 成立.

A5：(i) 对 A1(i) 中 $p \geqslant 2, \tilde{S}_n^{-2}\sum_{i=1}^{n} |\tilde{x}_i| \leqslant \alpha n^{-\frac{1}{2}}\log^{-3/p} n$.

(ii) 对 A4(ii) 中的 $0 < \theta < 1$ 及 A1(i) 中的 $p \geqslant 2$，

$$\tilde{S}_n^{-2}\sum_{i=1}^{n} |\tilde{x}_i| \leqslant \alpha n^{-1/2-\theta/p}\log^{-1} n.$$

A6：对 $t \in I$ 一致地有 $\max_{1\leqslant j\leqslant n} W_{nj}(t) \leqslant \alpha n^{-\frac{1}{2}}$.

**定理 2.3.3** 若假定 A1, A3 (i), (ii) 及 A4(i), A5(i) 与 A6 满足，则

$$\hat{\beta}_n \xrightarrow{a.s.} \beta. \tag{3.2.13}$$

若再满足 A3(iv)，则

$$\hat{g}_n(t) \to g(t), \forall t \in I. \tag{3.2.14}$$

**定理 2.3.4** 若条件 A1, A3(i), (ii) 与 A4(ii), A5(ii) 满足，则

$$E|\hat{\beta}_n - \beta|^p \to 0, \quad n \to \infty. \tag{3.2.15}$$

若再满足 A3(iv)，则

$$E|\hat{g}_n(t) - g(t)|^p \to 0, \quad n \to \infty, \forall t \in I. \tag{3.2.16}$$

### 3.2.2 定理的证明

**引理 3.2.1** 若随机变量 $|X| \leqslant 1$，则

$$E\exp(|X|) \leqslant \alpha\exp(E|X|^2).$$

证明：由不等式 $\exp(x)\leqslant 1+x+x^2\leqslant \exp(x+x^2)$，得

$$E\exp(|X|)\leqslant 1+E|X|+EX^2 \leqslant \exp(E|X|+EX^2)\leqslant \alpha\exp(EX^2).$$

于是引理 3.2.1 得证.

定理 3.2.1 的证明：容易看出

$$\begin{aligned}\tilde\beta_n-\beta &= \tilde S_n^{-2}\sum_{i=1}^n \tilde x_i(\tilde Z_{iG}-\tilde x_i\beta)\\ &= \tilde S_n^{-2}\sum_{i=1}^n \tilde x_i(g(t_i)-g_n(t_i))+\tilde S_n^{-2}\sum_{i=1}^n \tilde x_i e_i\\ &\overset{\mathrm{d}}{=} I_{1n}+I_{2n}.\end{aligned}\tag{3.2.17}$$

现往证

$$I_{1n}\xrightarrow{a.s.}0,\tag{3.2.18}$$

$$I_{2n}\xrightarrow{a.s.}0.\tag{3.2.19}$$

先证(3.2.18)，事实上

$$\begin{aligned}I_{1n} &\leqslant \max_{1\leqslant i\leqslant n}|g(t_i)-Eg_n(t_i)|\tilde S_n^{-2}\sum_{i=1}^n|\tilde x_i|\\ &\quad+\tilde S_n^{-2}\left|\sum_{i=1}^n \tilde x_i(Eg_n(t_i)-g_n(t_i))\right|\\ &\overset{\mathrm{d}}{=} I_{11n}+I_{12n}.\end{aligned}\tag{3.2.20}$$

使用胡舒合[102]中的(22)及 $g(t)$ 连续因此在 $t\in I$ 一致连续的性质，再由 A3(i)，(ii) 与(3.2.20)，我们有

$$\begin{aligned}&\max_{1\leqslant i\leqslant n}|g(t_i)-Eg_n(t_i)|\\ &\leqslant \alpha\left\{\max_{1\leqslant i\leqslant n}\left|\sum_{j=1}^n W_{nj}(t_i)-1\right|+\max_{1\leqslant i\leqslant n}\left|\sum_{j=1}^n W_{nj}(t_i)I[|t_j-t_i|>a]\right|\right.\\ &\quad\left.+\max_{1\leqslant i\leqslant n}\left|\sum_{j=1}^n W_{nj}(t_i)(g(t_j)-g(t_i))I[|t_j-t_i|\leqslant a]\right|\right\}\to 0,\end{aligned}\tag{3.2.21}$$

结合(3.2.20)对 $I_{11n}$ 的定义并应用条件 A2(i) 就证明了

$$I_{11n} \to 0, \quad n \to \infty. \tag{3.2.22}$$

参考 Erdös[30] 的截断技巧可证 $I_{12n} \xrightarrow{a.s.} 0$,事实上

$$I_{12n} = \sum_{j=1}^{n}\left(\tilde{S}_n^{-2}\sum_{i=1}^{n}\tilde{x}_i W_{nj}(t_i)\right)e_j \triangleq \sum_{j=1}^{n} a_{nj}e_j. \tag{3.2.23}$$

对任何取定的 $\varepsilon > 0$ 及某自然数

$$N_p > \frac{(1+p)\log n + 2\log\log n}{p\log\log n},$$

定义

$$e'_{nj} = e_j I[\,|a_{nj}e_j| \leqslant n^{-1}],\, I'_{12n} = \sum_{j=1}^{n} a_{nj}e'_{nj},$$

$$e''_{nj} = e_j I[n^{-1} < |a_{nj}e_j| \leqslant \varepsilon/N_p],\, I''_{12n} = \sum_{j=1}^{n} a_{nj}e''_{nj},$$

$$e'''_{nj} = e_j I[\,|a_{nj}e_j| > \varepsilon/N_p],\, I'''_{12n} = \sum_{j=1}^{n} a_{nj}e'''_{nj},\, \eta_n = \sqrt{\max_{1\leqslant j\leqslant n} Ee_j^2}.$$

由 Chebyshev 不等式,知对任何 $u > 0$,有

$$P(|I'_{12n}| > \varepsilon) \leqslant \exp\{-\varepsilon u/\eta_n\} E\exp\{u|I'_{12n}|/\eta_n\}. \tag{3.2.24}$$

若取 $u = \min\{\varepsilon/(2\eta_n\sum_{j=1}^{n}a_{nj}^2),\, n\eta_n\}$,则 $|ua_{nj}e'_{nj}|/\eta_n \leqslant 1$,于是应用引理 3.2.1 可得

$$\begin{aligned} P(|I'_{12n}| > \varepsilon) &\leqslant \exp\{-\varepsilon u/\eta_n\}\prod_{j=1}^{n} E\exp\{u|a_{nj}e'_{nj}|/\eta_n\} \\ &\leqslant \alpha\exp\left\{-\varepsilon u/\eta_n + u^2\sum_{j=1}^{n}a_{nj}^2\right\}. \end{aligned} \tag{3.2.25}$$

若 $\varepsilon/(2\eta_n\sum_{j=1}^{n}a_{nj}^2) > n\eta_n$,则 $u = n\eta_n$ 且 $u\sum_{j=1}^{n}a_{nj}^2 < \varepsilon/(2\eta_n)$,由此及(3.2.25)可得

$$P(|I'_{12n}| > \varepsilon) \leqslant \alpha\exp\{-\varepsilon n/2\}. \tag{3.2.26}$$

若 $\varepsilon/(2\eta_n\sum_{j=1}^{n}a_{nj}^2)\leqslant n\eta_n$，则 $u=\varepsilon/(2\eta_n\sum_{j=1}^{n}a_{nj}^2)$，于是

$$P(|I'_{12n}>\varepsilon)\leqslant\alpha\exp\left\{-\varepsilon^2/(4\eta_n\sum_{j=1}^{n}a_{nj}^2)\right\}.\quad(3.2.27)$$

由条件 A1 可推知 $\eta_n\leqslant\alpha$．而由条件 A2(i)，A3(iii) 可得

$$\sum_{j=1}^{n}a_{nj}^2\leqslant\sum_{j=1}^{n}(\max_{1\leqslant i\leqslant n}W_{nj}(t_i))^2\left(\tilde{S}_n^{-2}\sum_{i=1}^{n}|\tilde{x}_i|\right)^2\leqslant\alpha(\log n)^{-2}.\quad(3.2.28)$$

由此及(3.2.27)，或(3.2.26)，均可得

$$\sum_{n=1}^{\infty}P(|I'_{12n}|>\varepsilon)<\infty.\quad(3.2.29)$$

至于 $I''_{12n}$，我们有

$$\begin{aligned}P(|I''_{12n}|>\varepsilon)&\leqslant P(\text{至少存在 }1\leqslant i_1\leqslant i_2\leqslant\cdots\leqslant i_{N_p}\leqslant n\text{ 使得}\\&\qquad|a_{ni_1}e_{i_1}|>n^{-1},\cdots,|a_{ni_{N_p}}e_{i_{N_p}}|>n^{-1})\\&\leqslant\sum_{1\leqslant i_1<i_2<\cdots<i_{N_p}\leqslant n}\prod_{k=1}^{n}P(|e_{i_k}|>(n|a_{ni_k}|)^{-1})\\&\leqslant\sum_{1\leqslant i_1<i_2<\cdots<i_{N_p}\leqslant n}\prod_{k=1}^{n}n^p|a_{ni_k}|^p\mathrm{E}|e_{i_k}|^p.\quad(3.2.30)\end{aligned}$$

注意到条件 A1 蕴涵 $\max_{1\leqslant i\leqslant n}E|e_i|^p\leqslant\alpha$，而又由条件 A2(i) 与 A3(iii) 可推得 $\max_{1\leqslant j\leqslant n}|a_{nj}|\leqslant\alpha n^{-1/2}(\log n)^{-1}$．由此及(3.2.30) 知

$$\begin{aligned}\sum_{n=1}^{\infty}P(|I''_{12n}|>\varepsilon)&\leqslant\alpha\sum_{n=1}^{\infty}n^{N_p+p-pN_p/2}(\log n)^{-pN_p}\\&\leqslant\alpha\sum_{n=1}^{\infty}\frac{1}{n(\log n)^2}<\infty.\quad(3.2.31)\end{aligned}$$

注意应用切比雪夫不等式及条件 A1，A2(i) 与 A3(iii)，可得

$$\begin{aligned}P(|I'''_{12n}|>\varepsilon)&\leqslant P(\text{至少存在 }1\leqslant j\leqslant n\text{ 使得}|a_{nj}e_j|>\varepsilon/N_p)\\&\leqslant\alpha\sum_{j=1}^{n}|a_{nj}|^pE|e_j|^p\end{aligned}$$

$$\leqslant an(\max_{1\leqslant j\leqslant n}\max_{1\leqslant i\leqslant n}W_{nj}(t_i))^p(\tilde{S}_n^{-2}\sum_{i=1}^n\tilde{x}_i)^p$$

$$\leqslant an^{-1}(\log n)^{-p}, \tag{3.2.32}$$

于是由(3.2.32),即得

$$\sum_{n=1}^\infty P(|I'''_{12n}|>\varepsilon)\leqslant a\sum_{n=1}^\infty\frac{1}{n(\log n)^p}<\infty. \tag{3.2.33}$$

综合(3.2.29),(3.2.31) 与(3.2.32),我们得

$$\sum_{n=1}^\infty P(|I_{12n}|>\varepsilon)<\infty.$$

因此,由 Borel-Cantelli 引理即证明了

$$I_{12n}\xrightarrow{a.s.}0. \tag{3.2.34}$$

易见,(3.2.20),(3.2.22) 与(3.2.34) 一起就证明了(3.2.18).

仿效证(3.2.34) 的方法可证在条件 A1 及 A2(iii) 下, 有(3.2.19) 成立, 于是由(3.2.18), (3.2.19) 与(3.2.17) 即证得(3.2.22).

下面证(3.2.23). $\forall t\in I$

$$\begin{aligned}\tilde{g}_n(t)-g(t)&=(\tilde{g}_n(t)-g_n(t))+(g_n(t)-Eg_n(t))\\&\quad+(Eg_n(t)-g(t))\\&\triangleq II_{1n}+II_{2n}+II_{3n}.\end{aligned} \tag{3.2.35}$$

因 $II_{1n}=-\sum_{j=1}^n W_{nj}x_j(\tilde{\beta}_n-\beta)$,于是由 A3(iv) 及(3.2.22) 即得

$$II_{1n}\xrightarrow{a.s.}0. \tag{3.2.36}$$

又仿(3.2.34) 的证明可证在条件 A1 与 A3(iii) 下,有

$$II_{2n}=\sum_{j=1}^n W_{nj}(t)e_j\xrightarrow{a.s.}0,$$

类似于(3.2.21),我们有

$$II_{3n}\to 0, n\to\infty.$$

由(3.2.35),就得(3.2.23) 的证明. 至此定理 3.2.1 得证.

定理 3.2.2 的证明:由(3.2.17),得

$$|\tilde{\beta}_n-\beta|\leqslant\tilde{S}_n^{-2}\left|\sum_{i=1}^n\tilde{x}_i(g(t_i)-Eg_n(t_i))\right|$$

$$+\tilde{S}_n^{-2}\left|\sum_{i=1}^n\tilde{x}_i(Eg_n(t_i)-g_n(t_i))\right|+\tilde{S}_n^{-2}\left|\sum_{i=1}^n\tilde{x}_ie_i\right|.$$

于是由 $C_p$ 不等式,有

$$E|\tilde{\beta}_n-\beta|^p\leqslant\alpha(\tilde{S}_n^{-2p}\left|\sum_{i=1}^n\tilde{x}_i(g(t_i)-Eg_n(t_i)\right|^p$$

$$+\tilde{S}_n^{-2p}E\left|\sum_{i=1}^n\tilde{x}_i(Eg_n(t_i)-g_n(t_i))\right|^p$$

$$+\tilde{S}_n^{-2p}E\left|\sum_{i=1}^n\tilde{x}_ie_i\right|^p)$$

$$\stackrel{\mathrm{d}}{=}\alpha(E_{1n}+E_{2n}+E_{3n}).\tag{3.2.37}$$

由(3.2.20)与(3.2.22)知这里

$$E_{1n}\to0,n\to\infty.\tag{3.2.38}$$

根据 Whittle 不等式及条件 A2(ii),A3(iii),我们有

$$E_{2n}=\tilde{S}_n^{-2p}E\left|\sum_{i=1}^n[\tilde{x}_i(\sum_{j=1}^nW_{nj}(t_i)e_j)]\right|^p$$

$$\leqslant\alpha\tilde{S}_n^{-2p}[\sum_{j=1}^n(\sum_{i=1}^n\tilde{x}_iW_{nj}(t_i))^2(E|e_j|^p)^{2/p}]^{p/2}$$

$$\leqslant\alpha n^p\tilde{S}_n^{-p}(\max_{1\leqslant j\leqslant n}\max_{1\leqslant i\leqslant n}W_{nj}(t_i))^p\to0.\tag{3.2.39}$$

如同证(3.2.39),由 A2(ii) 可证

$$E_{3n}=\tilde{S}_n^{-2p}E\left|\sum_{i=1}^n\tilde{x}_ie_i\right|^p\leqslant\tilde{S}_n^{-2p}[\sum_{i=1}^n|\tilde{x}_i|^2(E|e_i|^p)^{2/p}]^{p/2}$$

$$\leqslant\alpha\tilde{S}_n^{-p}\to0.\tag{3.2.40}$$

综合(3.2.37)—(3.2.40),(3.2.9) 得证.

由(3.2.35),我们有

$$E|\tilde{g}_n(t)-g(t)|^p$$

$$\leqslant \alpha(E|\tilde{g}_n(t) - g_n(t)|^p + E|g_n(t) - Eg_n(t)|^p + |Eg_n(t) - g(t)|^p \tag{3.2.41}$$

利用条件 A3(iv) 与(3.2.9),得到

$$E|\tilde{g}_n(t) - g_n(t)|^p = \left|\sum_{j=1}^n W_{nj}(t)x_j\right|^p E|\tilde{\beta}_n - \beta|^p \to 0. \tag{3.2.42}$$

应用 Whittle 不等式及条件 A3(iii) 及 $\max\limits_{1\leqslant j\leqslant n} E|e_j|^p \leqslant \alpha$ 这一事实,我们有

$$E|g_n(t) - Eg_n(t)|^p \leqslant E\left|\sum_{j=1}^n W_{nj}(t)e_j\right|^p \leqslant \alpha\Big(\sum_{j=1}^n W_{nj}^2(t)\Big)^{p/2} \to 0. \tag{3.2.43}$$

在本定理的条件 A3(i),A3(ii) 下,并注意本文始终假定 $g(t)$ 在闭区间 $I$ 上连续,与(3.2.21) 同法可证

$$|Eg_n(t) - g(t)|^p \to 0, \quad n \to \infty, \forall t \in I. \tag{3.2.44}$$

因此,(3.2.41)—(3.2.44) 一起证明了(3.2.10).

**引理 3.2.2** 对任何 $r > 0$

$$E\{(1 + N^+(Z_i))^{-r} | (Z_i,\delta_i)\} \leqslant \alpha n^{-r}\{\overline{H}_S(Z_i) - n^{-1}\}^{-r},$$

$$E\{(1 + N^+(Z_i))^{-r} | (Z_i,\delta_i),(Z_j,\delta_j)\} \leqslant \alpha n^{-r}\{\overline{H}_S(Z_i) - 2n^{-1}\}^{-r}, \quad 1 \leqslant i,j \leqslant n,$$

其中 $\overline{H}_S(\cdot) = \frac{1}{n}\sum_{i=1}^n \overline{H}_{iS}(\cdot)$.

证明:见 Koul,Susarlar 与 Ryzin[47].

定理 3.2.3 的证明:容易看出

$$\hat{\beta}_n - \beta = \tilde{S}_n^{-2}\sum_{i=1}^n \tilde{X}_i(Z_{i\hat{G}_n} - Z_{iG})I[Z_i \leqslant M_n] - \tilde{S}_n^{-2}\sum_{i=1}^n \tilde{x}_i\sum_{j=1}^n W_{nj}(t_i)(Z_{j\hat{G}_n} - Z_{jG})I[Z_j \leqslant M_n]$$

$$+\tilde{S}_n^{-2}\sum_{i=1}^n\tilde{x}_i(Z_{iG}I[Z_i\leqslant M_n]-Z_{iG})$$

$$+\tilde{S}_n^{-2}\sum_{i=1}^n\tilde{x}_i\sum_{j=1}^n W_{nj}(t_i)(Z_{jG}-Z_{jG}I[Z_j\leqslant M_n])$$

$$+\tilde{S}_n^{-2}\sum_{i=1}^n\tilde{x}_i(\tilde{Z}_{iG}-\tilde{x}_i\beta)\overset{\mathrm{d}}{=}\sum_{i=1}^5\Delta_{ni}. \tag{3.2.45}$$

首先证

$$\Delta_{n1}\xrightarrow{a.s.}0. \tag{3.2.46}$$

由前面 $Z_{i\hat{G}_n}$ 及 $Z_{iG}$ 的定义,并注意应用 $C_p$ 不等式,得

$$E|\Delta_{n1}|^p=\tilde{S}_n^{-2p}n^{p-1}\sum_{i=1}^n|x_i|^pE\{\delta_i|Z_i|^pI[Z_i\leqslant M_n]$$

$$E[|\hat{G}_{nS}^{-1}(Z_i)-G_S^{-1}(Z_i)||(Z_i,\delta_i)]\}. \tag{3.2.47}$$

注意到 $\hat{G}_{nS}^{-1}$ 与 Susarla 与 Ryzin[74] 中 $W_n$ 相同,于是我们有

$$\hat{G}_{nS}^{-1}(Z_i)\leqslant(n+1)/(N^+(Z_i)+1). \tag{3.2.48}$$

由(3.2.48)与不等式:$|x-y|\leqslant|\log x-\log y|$, $0<x\leqslant1,0<y\leqslant1$,我们有

$$E_i(\hat{G}_{nS},G_S)$$

$$\overset{\mathrm{d}}{=}E[|\hat{G}_{nS}^{-1}(Z_i)-G_S(Z_i)|^p|(Z_i,\delta_i)]$$

$$=E[\hat{G}_{nS}^{-p}(Z_i)G_S^{-p}(Z_i)|\hat{G}_{nS}(Z_i)-G_S(Z_i)|^p|(Z_i,\delta_i)]$$

$$\leqslant G_S^{-p}(Z_i)E^{1/2}\left[\left(\frac{N^+(Z_i)+1}{n+1}\right)^{-2p}\Bigg|(Z_i,\delta_i)\right]$$

$$\times E^{1/2}[|\log\hat{G}_{nS}(Z_i)-\log G_S(Z_i)|^{2p}|(Z_i,\delta_i)]. \tag{3.2.49}$$

定义 $\beta_j(z)=I[\delta_j=0,Z_j\leqslant z]$,则

$$\log\hat{G}_{nS}(Z_i)-\log G_S(Z_i)$$

$$=\left[-\frac{1}{n}\sum_{j=1}^n\beta_j(Z_i)\overline{H}_S(Z_j)-\log G_S(Z_i)\right]$$

$$+\left[-\sum_{j=1}^n\beta_j(Z_i)\sum_{l=2}^n\frac{1}{l}(2+N^+(Z_j))^{-l}\right]$$

$$+\left[-\frac{1}{n}\sum_{j=1}^{n}\beta_j(Z_i)\left(\frac{n}{2+N^+(Z_j)}-\overline{H}_S^{-1}(Z_j)\right)\right]$$

$$\stackrel{\mathrm{d}}{=} T_{n1}+T_{n2}+T_{n3}. \tag{3.2.50}$$

既然

$$-E[\beta_j(Z_i)\overline{H}_S^{-1}(Z_i)\,|\,(Z_i,\delta_i)]=-\int_0^{Z_i}\frac{F_{jS}(u)}{\overline{H}_S(u)}\mathrm{d}G(u),$$

我们有

$$\frac{1}{n}\sum_{j=1}^{n}E[-\beta_j(Z_i)\overline{H}_S(Z_i)\,|\,(Z_i,\delta_i)]$$

$$=-\frac{1}{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\int_0^{Z_i}\frac{F_{jS}(u)}{\overline{H}_S(u)}\mathrm{d}G(u)-\frac{1}{n}I[\delta_i=0]\overline{H}_S^{-1}(Z_i)$$

$$=\log G_S(t)+\frac{1}{n}\int_0^{Z_i}\frac{F_{iS}(u)}{\overline{H}_S(u)}\mathrm{d}G(u)$$

$$-\frac{1}{n}I[\delta_i=0]\overline{H}_S^{-1}(Z_i), \tag{3.2.51}$$

由(3.2.50)中 $T_{n1}$ 的定义,由(3.2.51),得

$$T_{n1}=\frac{1}{n}\sum_{j=1}^{n}\{-\beta_j(Z_i)\overline{H}_S^{-1}(Z_j)-E[-\beta_j(Z_i)\overline{H}_S^{-1}(Z_i)\,|\,(Z_i,\delta_i)\}$$

$$+\frac{1}{n}\int_0^{Z_i}\frac{F_{iS}(u)}{\overline{H}_S(u)}\mathrm{d}G(u)-\frac{1}{n}I[\delta_i=0]\overline{H}_S^{-1}(Z_i)$$

$$\stackrel{\mathrm{d}}{=} T_{n11}+T_{n12}+T_{n13}. \tag{3.2.52}$$

利用 Dharmadhikar-Jogdeo(D-J) 不等式,可得

$$E\{|T_{n11}|^{2p}\,|\,(Z_i,\delta_i)\}$$

$$\leqslant an^{-2p}E\left\{\left|\sum_{j=1}^{n}[\beta_j(Z_i)\overline{H}_S^{-1}(Z_j)\right.\right.$$

$$\left.-E(\beta_j(Z_i)\overline{H}_S^{-1}(Z_j)\,|\,(Z_i,\delta_i)]\right|^{2p}|\,(Z_i,\delta_i)\}$$

$$\leqslant an^{-2p}(n-1)^{p-1}\sum_{\substack{j=1\\ j\neq i}}^{n}E\,|\beta_j(Z_i)\overline{H}_S^{-1}(Z_j)|^{2p}\,|\,(Z_i,\delta_i)\}$$

$$\leqslant \alpha n^{-p-1}\sum_{\substack{j=1\\ j\neq n}}^{n}\int_0^{Z_i}\frac{F_{jS}(u)}{\overline{H}_S^{2p}(u)}\mathrm{d}G(u)$$

$$\leqslant \alpha n^{-p}\overline{H}_S^{-2p}(Z_i). \tag{3.2.53}$$

又易见

$$E\{|T_{n12}|^{2p}|(Z_i,\delta_i)\}\leqslant \alpha n^{-2p}\overline{H}_S^{-2p}(Z_i), \tag{3.2.54}$$

$$E\{|T_{n13}|^{2p}|(Z_i,\delta_i)\}\leqslant \alpha n^{-2p}I[\delta_i=0]\overline{H}_S^{-2p}(Z_i). \tag{3.2.55}$$

由(3.2.52)与(3.2.55),我们有

$$E\{|T_{n1}|^{2p}|(Z_i,\delta_i)\}\leqslant \alpha n^{-p}\overline{H}_S^{-2p}(Z_i), \tag{3.2.56}$$

又因

$$\sum_{l=2}^{\infty}\frac{1}{l}(2+N^+(Z_j))^{-l}\leqslant\frac{3}{2}\frac{1}{(1+N^+(Z_j))^2},$$

于是我们有

$$E\{|T_{n2}(Z_i)|^{2p}|(Z_i,\delta_i)\}$$

$$\leqslant \alpha E\left\{\left|\sum_{j=1}^{n}\beta_j(Z_i)(1+N^+(Z_j))^{-2}\right|^{2p}\Big|(Z_i,\delta_i)\right\}$$

$$\leqslant \alpha n^{2p-1}\sum_{j=1}^{n}E\{\beta_j(Z_i)(1+N^+(Z_j))^{-4p}\Big|(Z_i,\delta_i)\}$$

$$\leqslant \alpha n^{2p-1}\sum_{j=1}^{n}E\{\beta_j(Z_i)E[(1+N^+(Z_j))^{-4p}|$$

$$(Z_i,\delta_i),(Z_j,\delta_j)]|(Z_i,\delta_i)\}. \tag{3.2.57}$$

在(3.2.57)中应用引理2.3.2的第二个不等式,知 $n$ 充分大时

$$E\{|T_{n2}(Z_i)|^{2p}|(Z_i,\delta_i)\}$$

$$\leqslant \alpha n^{-2p-1}\sum_{j=1}^{n}E\{\beta_j(Z_i)[\overline{H}_S(Z_j)-n^{-1}]^{-4p}|(Z_i,\delta_i)\}$$

$$\leqslant \alpha n^{-2p-1}\sum_{j=1}^{n}E[2^{4p}\beta_j(Z_i)\overline{H}_S^{-4p}(Z_i)|(Z_i,\delta_i)\}$$

$$\leqslant \alpha n^{-2p}\overline{H}_S^{-4p}(Z_i). \tag{3.2.58}$$

通过简单的代数运算,可得

$$T_{n3}\leqslant \frac{1}{n}\sum_{j=1}^{n}\beta_j(Z_i)\frac{n}{(1+N^+(Z_j))^2}$$

$$-\frac{1}{n}\sum_{j=1}^{n}\beta_j(Z_i)(n/(1+N^+(Z_j))-\overline{H}^{-1}(Z_j))$$

$$\stackrel{\mathrm{d}}{=} T_{n31}-T_{n32}. \tag{3.2.59}$$

由(3.2.57),(3.2.59),我们有

$$E\{|T_{n31}|^{2p}|(Z_i,\delta_i)\}\leqslant \alpha n^{-2p}\overline{H}_S^{-4p}(Z_i). \tag{3.2.60}$$

使用 $C_p$ 不等式,得

$$E\{|T_{n32}|^{2p}|(Z_i,\delta_i)\}$$

$$\leqslant \alpha n^{-1}\sum_{j=1}^{n}E\{\beta_j(Z_i)E\left[\left|\frac{n}{1+N^+(Z_i)}-\frac{1}{\overline{H}_S(Z_j)}\right|^{2p}\right.$$

$$\left.\Big|(Z_i,\delta_i),(Z_j,\delta_j)\right]|(Z_i,\delta_i)\}. \tag{3.2.61}$$

由 Schwarz 不等式,我们有

$$E\left[\left|\frac{n}{1+N^+(Z_j)}-\frac{1}{\overline{H}_S(Z_j)}\right|^{2p}\Big|(Z_i,\delta_i),(Z_j,\delta_j)\right]$$

$$\leqslant \overline{H}_S^{-2p}(Z_j)E^{1/2}[(1+N^+(Z_j))^{-4p}|(Z_i,\delta_i),(Z_j,\delta_j)]$$

$$\times E^{1/2}[|(1+N^+(Z_j))-n\overline{H}_S(Z_j)|^{4p}|(Z_i,\delta_i),(Z_j,\delta_j)]. \tag{3.2.62}$$

记 $\overline{H}_S^{(-i)}(Z_j)=\frac{1}{n-1}\sum_{\substack{l=1\\ l\neq i}}^{n}H_{lS}(Z_j)$,于是由 D-J 不等式,我们有

$$E[|(1+N^+(Z_j))-n\overline{H}_S(Z_j)|^{-4p}|(Z_i,\delta_i),(Z_j,\delta_j)]$$

$$\leqslant E\left[\left|\sum_{\substack{l=1\\ l\neq i}}^{n}I[Z_l>Z_j]-(n-1)\overline{H}_S^{(-i)}(Z_j)\right|^{4p}|(Z_i,\delta_i),(Z_j,\delta_j)\right]$$

$$+1+I[Z_i>Z_j]+H_{iS}^{4p}(Z_j)$$

$$\leqslant \alpha(n-1)^{2p-1}\sum_{\substack{l=1\\ l\neq i}}^{n}[H_{lS}+H_{lS}^{4p}(Z_j)]+3\leqslant \alpha n^{2p}. \tag{3.2.63}$$

根据(3.2.61)—(3.2.63)与引理3.2.2,我们得

$$
\begin{aligned}
&E\{|T_{n32}|^{2p}|(Z_i,\delta_i)\}\\
&\leqslant \alpha n^{-1}\sum_{j=1}^{n}E\{\beta_j(Z_i)\overline{H}_S^{-2p}(Z_j)n^{-p}[\overline{H}_S(Z_j)-2n^{-1}]^{-2p}|(Z_i,\delta_i)\}\\
&\leqslant \alpha n^{-p-1}\sum_{j=1}^{n}E\{\beta_j(Z_i)2^{2p}\overline{H}_S^{-4p}(Z_j)|(Z_i,\delta_i)\}\\
&\leqslant \alpha n^{-p}\overline{H}_S^{-4p}(Z_i).
\end{aligned}\tag{3.2.64}
$$

于是,(3.2.59),(3.2.60)与(3.2.64)一起就给出了

$$E\{|T_{n3}|^{2p}|(Z_i,\delta_i)\}\leqslant \alpha n^{-p}\overline{H}_S^{-4p}(Z_i).\tag{3.2.65}$$

由(3.2.50),(3.2.56),(3.2.58)与(3.2.65),得

$$E\{|\log\hat{G}_{nS}(Z_i)-\log G_S(Z_i)|^{2p}|(Z_i,\delta_i)\}\leqslant \alpha n^{-p}\overline{H}_S^{-4p}(Z_i).\tag{3.2.66}$$

结合(3.2.66),(3.2.49)并应用引理3.2.2,即得

$$
\begin{aligned}
&E_i(\hat{G}_{nS},G_S)\\
&\leqslant \alpha G_S^{-p}(Z_i)((n+1)/n)^p\{\overline{H}_S(Z_i)-n^{-1}\}^{-p}n^{-p/2}\overline{H}_S^{-2p}(Z_i)\\
&\leqslant \alpha n^{-p/2}G_S^{-p}(Z_i)\overline{H}_S^{-3p}(Z_i).
\end{aligned}\tag{3.2.67}
$$

再由(3.2.47),(3.2.48)与(3.2.67),有

$$
\begin{aligned}
&E|\Delta_{n1}|^p\\
&\leqslant \alpha\tilde{S}_n^{-2p}n^{p/2-1}\sum_{i=1}^{n}|\tilde{x}_i|^p\\
&\quad\cdot E\{\delta_i|Z_i|^pI[Z_i\leqslant M_n]G_S^{-p}(Z_i)\overline{H}_S^{-3p}(Z_i)\}\\
&\leqslant \alpha\tilde{S}_n^{-2p}n^{p/2-1}\sum_{i=1}^{n}|\tilde{x}_i|^p\int_{-\infty}^{M_n}|u|G_S^{-p}(u)\overline{H}_S^{-3p}(u)G_S(u)\mathrm{d}F_i(u)\\
&\leqslant \alpha\tilde{S}_n^{-2p}n^{p/2-1}\sum_{i=1}^{n}|\tilde{x}_i|^p\overline{H}_S^{-3p}(M_n)\cdot\max_{1\leqslant i\leqslant n}E\{|Y_i|^pG_S^{1-p}(Y_i)\}\\
&\leqslant \alpha[n^{1/2-1/p}\log^{1/p}n\tilde{S}_n^{-2}(\sum_{i=1}^{n}|x_i|^p)^{1/p}]^p
\end{aligned}
$$

$$\leqslant \alpha(n^{1/2-1/p}\log^{1/p}n\tilde{S}_n^{-2}\sum_{i=1}^{n}|\tilde{x}_i|)^p \leqslant \frac{\alpha}{n\log^2 n} \qquad (3.2.68)$$

(注意在(3.2.68)的最后一个不等式中利用到条件 A5(i)). 我们再利用 Tchebyshev 不等式及(3.2.68),有

$$\sum_{n=1}^{\infty}P(|\Delta_{n1}|>\varepsilon)\leqslant\sum_{n=1}^{\infty}\frac{\alpha}{n\log^2 n}\leqslant\infty. \qquad (3.2.69)$$

于是由 Borel-Cantelli 引理即证得(3.2.46).

完全套用证(3.2.46)的方法,可证在条件 A6,A5(i) 下,有

$$\Delta_{n2}\xrightarrow{a.s.}0. \qquad (3.2.70)$$

又易见

$$\begin{aligned}\Delta_{n3} &= -\tilde{S}_n^{-2}\sum_{i=1}^{n}\tilde{x}_i\{\delta_iZ_iG_S^{-1}(Z_i)I[Z_i>M_n] \\ &\quad - E\delta_iZ_iG_S^{-1}(Z_i)I[Z_i>M_n]\} \\ &\quad - \tilde{S}_n^{-2}\sum_{i=1}^{n}\tilde{x}_iE\delta_iZ_iG_S^{-1}(Z_i)I[Z_i>M_n] \\ &\stackrel{d}{=} \Delta_{n31}+\Delta_{n32}.\end{aligned} \qquad (3.2.71)$$

由 Chebyshev 不等式与 Whittle 不等式,得

$$\begin{aligned}&P(|\Delta_{n31}|>\varepsilon) \\ &\leqslant \alpha E|\tilde{S}_n^{-2}\sum_{i=1}^{n}\tilde{x}_i(\delta_iZ_iG_S^{-1}(Z_i)I[Z_i>M_n] \\ &\quad - E\delta_iZ_iG_S^{-1}(Z_i)I[Z_i>M_n])|^p \\ &\leqslant \alpha\tilde{S}_n^{-2p}\left\{\sum_{i=1}^{n}\tilde{x}_i^2E^{2/p}[\delta_i|Z_i|G_S^{-1}(Z_i)I[Z_i>M_n]]^p\right\}^{p/2} \\ &\leqslant \alpha\max_{1\leqslant i\leqslant n}E[|Y_i|^pG_S^{1-p}(Y_i)]\tilde{S}_n^{-2p}n^{p/2-1}(\sum_{i=1}^{n}|\tilde{x}_i|)^p \\ &\leqslant \alpha\left(\tilde{S}_n^{-2}n^{1/2-1/p}\sum_{i=1}^{n}|\tilde{x}_i|\right)^p.\end{aligned} \qquad (3.2.72)$$

利用条件 A5(i) 与(3.2.72),即得

$$\sum_{n=1}^{\infty} P(|\Delta_{n31}| > \epsilon) \leqslant \alpha \sum_{n=1}^{\infty} \frac{1}{n(\log n)^3} < \infty. \quad (3.2.73)$$

于是由 Borel-Cantelli 引理知

$$\Delta_{n31} \xrightarrow{a.s.} 0. \quad (3.2.74)$$

既然 $E(\delta_i Z_i G_S^{-1}(Z_i) I[Z_i > M_n]) = \int_{M_n}^{\tau_{F_i}} s \mathrm{d}F_i(s)$，及条件 A1 蕴涵

$$\max_{1 \leqslant i \leqslant n} \int_{M_n}^{\tau_{F_i}} s \mathrm{d}F_i(s) \to 0, \quad n \to \infty,$$

因而结合条件 A5(i)，即得

$$\Delta_{n32} \to 0, \quad n \to \infty. \quad (3.2.75)$$

于是，(3.2.71)，(3.2.74) 与(3.2.75) 一起就证明了

$$\Delta_{n3} \xrightarrow{a.s.} 0, \quad n \to \infty. \quad (3.2.76)$$

套用证(3.2.76) 的方法，可证在 A6，A5(i) 下，有

$$\Delta_{n4} \xrightarrow{a.s.} 0. \quad (3.2.77)$$

注意到

$$|\Delta_{n5}| = |\tilde{\beta}_n - \beta| \leqslant I_{11n} + I_{12n} + I_{2n}, \quad (3.2.78)$$

其中 $I_{11n}, I_{12n}$ 与 $I_{2n}$ 如同(3.2.20)，(3.2.17) 所定义，由定理 3.2.1 的证明知，在假定 A3(i)，(ii) 与 A5(i)(A5(i) 蕴涵条件 A2(i)) 下，(2.3.23) 在此亦成立.

欲证这里

$$I_{12n} = \sum_{j=1}^{n} a_{nj} e_j \xrightarrow{a.s.} 0, \quad (3.2.79)$$

我们可使用(3.2.34) 的证明方法. 根据它的证明，我们仅需要验证 $a_{nj}, 1 \leqslant j \leqslant n$ 满足

a) $\sum_{j=1}^{n} a_{nj}^2 = o(\log^{-1} n)$,

b) $\max_{1 \leqslant j \leqslant n} |a_{nj}| \leqslant \alpha n^{-1/p} \delta_n$,

$$\text{c)} \sum_{n=1}^{\infty}\sum_{j=1}^{n}|a_{nj}|^{p}<\infty .$$

由 $a_{nj}$ 的定义，并注意应用 A6 与 A5(i)，可得

$$\begin{aligned}
&\sum_{n=1}^{\infty}\Big(\sum_{j=1}^{n}a_{nj}^{2}\Big)^{p/2}\\
&\leqslant \sum_{n=1}^{\infty}n^{p/2-1}\sum_{j=1}^{n}|a_{nj}|^{p}\\
&\leqslant \alpha\sum_{n=1}^{\infty}\max_{1\leqslant j\leqslant n}\max_{1\leqslant i\leqslant n}W_{nj}^{p}(t_{i})\cdot n\cdot n^{-p/2}\log^{-3}n\cdot n^{p/2-1}\\
&\leqslant \alpha\sum_{i=1}^{n}\frac{1}{n^{p/2}\log^{3}n}<\infty .
\end{aligned} \tag{3.2.80}$$

从上面第二个不等式到最后不等式可看出 c) 在此成立，再次利用 (2.3.80) 及级数的收敛性，知存在 $\delta_n\to 0, n\to\infty$ 使得

$$\Big(\sum_{j=1}^{n}a_{nj}^{2}\Big)^{p/2}\leqslant \alpha n^{-1}\delta_{n}^{p}. \tag{3.2.81}$$

由此即得 $\sum_{j=1}^{n}a_{nj}^{2}\leqslant \alpha n^{-2/p}\delta_{n}{}^{2}$，且 $\max_{1\leqslant j\leqslant n}|a_{nj}|\leqslant\Big(\sum_{j=1}^{n}a_{nj}^{2}\Big)^{1/2}$ $\leqslant \alpha n^{-1/p}\delta_{n}$. 这说明在此 a), b) 被验证，因而(3.2.79) 得证.

记 $a_{nj}'=\tilde{x}_{j}/\tilde{S}_{n}^{2}$. 于是利用条件 A5(i) 可验证 $a_{nj}', j=1,2,\cdots,n$ 满足类似于 $a_{nj}$ 所满足的条件 a), b), c)，因而同上面一样的道理有

$$I_{2n}=\sum_{j=1}^{n}a_{nj}'e_{j}\xrightarrow{a.s.}0. \tag{3.2.82}$$

因而(3.2.22)，(3.2.78)，(3.2.79) 与(3.2.82) 一起证明了

$$\Delta_{n5}\xrightarrow{a.s.}0. \tag{3.2.83}$$

综合 (3.2.45)，(3.2.46)，(3.2.70)，(3.2.76)，(3.2.77) 与 (3.2.83)，即知(3.2.13) 得证.

类似于证(3.2.8) 的证法可证(3.2.14)，所不同的是在证 (3.2.8) 时利用 (3.2.7) 的地方，在证 (2.3.14) 时却利用了

(3.2.13). 至此,定理 3.2.3 得证.

定理 3.2.4 的证明:由(3.2.45),得

$$E|\hat{\beta}_n-\beta|^p \leqslant \alpha \sum_{i=1}^{5} E|\Delta_{ni}|^p. \qquad (3.2.84)$$

仿(3.2.78) 的证明,在条件 A4(ii),A5(ii) 下,有

$$E|\Delta_{n1}|^p \leqslant \alpha n^{p/2-1+\theta}\Big[\tilde{S}_n^{-2}\Big(\sum_{i=1}^{n}|\tilde{x}_i|^p\Big)^{1/p}\Big]^p$$

$$\leqslant \alpha n^{p/2-1+\theta}\Big(\tilde{S}_n^{-2}\sum_{i=1}^{n}|\tilde{x}_i|\Big)^p \to 0. \qquad (3.2.85)$$

注意到 $0\leqslant W_{nj}(t)\leqslant 1, \forall\, t\in I$,因此再次利用 A4(ii),A5(ii),可得

$$E|\Delta_{n2}|^p \leqslant \alpha n^{p/2-1+\theta}\tilde{S}_n^{-2p}\sum_{j=1}^{n}\Big|\sum_{i=1}^{n}\tilde{x}_i W_{nj}(t_i)\Big|^p$$

$$\leqslant \alpha n^{p/2+\theta}\max_{1\leqslant j\leqslant n}\max_{1\leqslant i\leqslant n}W_{ni}^p(t_i)\Big(\tilde{S}_n^{-2}\sum_{j=1}^{n}|\tilde{x}_i|\Big)^p$$

$$\to 0. \qquad (3.2.86)$$

由 Whittle 不等式,条件 A1 与 A5(ii),我们可得

$$E|\Delta_{n4}|^p \leqslant \alpha\tilde{S}_n^{-2p}\Big\{\sum_{j=1}^{n}\Big(\sum_{i=1}^{n}\tilde{x}_i W_{nj}(t_i)\Big)^2 E^{2/p}\big[|Y_j|^p G_S^{1-p}(Y_j)\big]\Big\}^{p/2}$$

$$\leqslant \alpha\tilde{S}_n^{-2p}\cdot n^{p/2-1}\sum_{j=1}^{n}\Big|\sum_{i=1}^{n}\tilde{x}_i W_{nj}(t_i)\Big|^p \to 0, \qquad (3.2.87)$$

$$E|\Delta_{n3}|^p \leqslant \tilde{S}_n^{-2p}\Big(\sum_{i=1}^{n}\tilde{x}_i^2 E^{2/p}\big[|Y_i|G_S^{1-p}(Y_i)\big]\Big)^{p/2}$$

$$\leqslant \alpha n^{p/2-1}\Big(\tilde{S}_n^{-2}\sum_{i=1}^{n}|\tilde{x}_i|\Big)^p \to 0. \qquad (3.2.88)$$

既然 $\tilde{S}_n^{-1} = \tilde{S}_n^{-2}\Big(\sum_{i=1}^{n}\tilde{x}_i^2\Big)^{1/2} \leqslant \tilde{S}_n^{-2}\sum_{i=1}^{n}|\tilde{x}_i|$,于是由 A5(ii) 知(3.2.40)在此成立.对(3.2.37)所定义的 $E_{2n}$,由(3.2.39)并利用本定理中条件 A5(ii),与证(2.3.86)一样的道理可得

$$E_{2n} \leqslant \alpha \tilde{S}_n^{-2p} \Big[ \sum_{j=1}^{n} \Big( \sum_{i=1}^{n} \tilde{x}_i W_{nj}(t_i) \Big)^2 \Big]^{p/2}$$

$$\leqslant \alpha n^{p/2} \Big( \tilde{S}_n^{-2} \sum_{i=1}^{n} |\tilde{x}_i| \Big)^p \max_{1 \leqslant j \leqslant n} \max_{1 \leqslant i \leqslant n} W_{nj}{}^p(t_i)$$

$$\to 0. \tag{3.2.89}$$

又(3.2.38)在此成立,由此及(3.2.37),(3.2.40)与(3.2.89)知在此有

$$E|\Delta_{n5}|^p = E|\tilde{\beta}_n - \beta|^p \to 0, \quad n \to \infty. \tag{3.2.90}$$

因而由(3.2.84)—(3.2.88),(3.2.90)即得(3.2.15)的证明.

(3.2.16)的证明完全类似于(3.2.10)的证明,只是在证(3.2.10)时利用了(3.2.9)的地方,在证(3.2.16)时利用了(3.2.15). 至此,定理3.2.4证毕.

# 参考文献


Amemiya J. Advanced Economics. Cambridge: Harvard University Press, 1985

Andertsen P. Counting Process Models For Life History Data: A Review. Scand J. Statist., 1985, 12: 97～185

Benedetti J K. On the nonparametric estimation ofregression function. J. R. Statist. Soc., B, 1977, 39: 248～253

Berger J. Nonparametric Regression Analysis of Longitudinal Data. New York: Springer-Verlag, 1988

Bickel P J, Götze Z, Van Zwet W R. The Edgeworth expansion for U-statistics of degree two. Ann. Statist., 1986, 14: 1463～1481

Birgé L. Estimating a density under order restrictions: Nonasympototic minimax risk, Ann. Statist., 1987, 15(3): 995～1012

Blum J R, Susarla V. Maximal Deviation Theory of Density and FailureRate Function Estimates Based on Censored Data. Multivariate Analysis V (P. R. Krishnaich, Ed). Amsterdam: North-Holland, 1980. 213～222


Bickel P J, Götze Z, Van Zwet W R. The Edgeworth expansion for U-statistics of degree two. Ann. Statist., 1986, 14: 1463～1481


Bresley N E, Crowley J. A large sample study of life table and product-limit estimates under random censorship. Ann. Statist., 1974, 6: 437～457

Buckley J, James I. Linear regression with censored data. Biometrika, 1979, 66: 429～436

Burke M D, Csörgo S, Horváth L. Strong approximation of some biometric

estimates under random censorship. Z. Wahrsch. Verw. Gebiete, 1981, 56:87 ～112

Burke M D, Csörgo S, Horváth L. A correction to an improvement of "strong approximations of some biometricestimates under random censorship", Probab. Th. Rel. Fields, 1988, 79:51～57

Callaert H, Janssen P. The Berry-Essen theorem for U-statistics, Ann. Statist., 1978, 6(2):417～421

Campbell G. Nonparametric bivariate estimators with censored data, Biometrika, 1981, 68:417～422

Campbell G, Földes A. Asympototic properties of several nonparametric multivariate distribution function estimates under random censoring, In Survival Analysis (Crowley, J. and Johnson, R. A., Eds), IMS lecture notes, 1982, 2: 243～251

Chang M N, Rao P V. Berry-Essen bound for the Kaplan-Meier estimator. Commun. Statist. -Theory. Meth., 1989, 18(12):4647～4664

Chang M N. Edgeworth expansion for the Kaplan-Meier estimator. Commun. Statist.-Theory Meth. 1991, 20:2479～2494

Chang M N . Moments of the Kaplan-Meier estimator, Sankhyā: The Indian Journal of Statistics, Series A, 1991, 53:27～50

Chao M T, Lo S H. Some representations of the nonparametric maximum likelihood estimators with truncated data, Ann. Statist., 1988, 16(2): 661～668

Chen H. Convergence rates for parameter components in a partly linear models. Ann Statist, 1988, 16(1):136～146

Chorai J K, Pattanaik L M. $L_1$-consistency of the kernel density estimators based on randomly right censored data, Commun. Statist. —Theory. Meth., 1990, 19 (8):2853～2870

Chow Y S, Henry Teicher. Probability Theory. New York: Spring-Verlag, 1978

Csörgö S, Horváth L. The rate of strong uniform consistency for the product-limit estimator. Z. Wahrscheinlichkeitstheorie verw. Gebiete. 1983, 62:411～426

Csörgö M, CombayE, Horváth L.. Central limit theorems for $L_p$-distances of kernel estimators of densities under random censorship, Ann. Statist., 1991, 19

(4):1813～1831

Dabrowska D M. Uniform consistency of the kernel conditional Kaplan-Meier estimator. 1989,17(3):1157～1167

Diehl S, Stute W. Kernel density and hazard function estimation in the presence of censoring, J. Mult, Analy., 1988,25:299～310

Efron B. Bootstrap methods: another look at the jackknife. Ann. Statist., 1979, 7:1～26

Efron B. Censored data and the bootstrap. J. Amer. Stat. Assoc., 1981,76:312～321

Engle R, Granger C, Rice J *et al*. Nonparametric estimates of the relation between weather and electricity sales, J. Amer. Statist. Assoc., 1986,81:310～320

Erdos P. On a theorem of Hsu and Robbins. Ann. Math. Statist., 1949,20:286～296

Fleming T R, Harrington D P. Counting processes and survival analysis. New York: John Wiley & Sons, Inc, 1991

Földes A, Rejtö L. A LIL type result for product limit estimator, Z. Wahrsch. Verw, Gebiete. 1981,56:75～86

Földes A, Rejtö L. Strong uniformly consistency for nonparametric survival estimators from randomly censored data, Ann. Statist., 1981,9:122～129

Földes A, Rejtö L., Winter B B. Strong consistency properties of nonparametric estimation for randomly censored data, I: The product-limit estimator. Perioda Math. Hung, 1980,11:233～250

Földes A, Rejtö L., Winter B B. Strong consistency properties of nonparametric estimation for randomly censored data, II: Estimation of density and failure rate. Perioda Math. Hung, 1981,12:15～29

Gijbels Irène, Veraverbeke Noël. Almost sure asympototic representation for a class of functionals of the Kaplan-Meier estimator. Ann Statist, 1991,19(3):1457～1470

Gijbels Ireéne, Wang J L. Strong representation of the survival function estimator for truncated and censored data with application, J. Multi. Analy., 1993,47(2):210～229

Gill R D. Censoring and stochastic integrals. Amsterdam: Mathematical Centre

Tract, 1980

Gill R D. Large sample behavior of the product-limit estimator on the whole line. Ann Statist, 1983, 11:49～58

Gu M G. Functional laws of the iterated logarithm for the product limit estimator of a distribution function under random censorship or truncation, Ann. Proba., 1990, 18(1):160～189

Hall P, Heyde C C. Martingale Limit Theory And Its Applications. New York: Academic Press, 1980

Horváth L. Dropping continuity and independence assumptions in random censorship models. Studia Sci. Math. Hungar, 1983, 15:381～389

Jackson J C. The analysis of quasar samples. Monthly Notices Roy, Astronom. Soc., 1974, 168:281～295

James I R, Smith P J. Consistency results for linear regression with censored data, Ann. Statist., 1984, 12(2):590～600

Kaplan E L, Meier P. Nonparametric estimation from imcomplete observations. J. Amer, Statist. Assoc., 1958, 53:457～481

Kiefer J. On the Large Deviation of Empirical Distribution of Vectors chance Variables and a Law of the Integrated Logarithm, Pacific J. Math., 1961, 11: 649～660

Koul H, Susarlar V, Van Ryzin J. Regression analysis with randomly right-censored data. Ann Statist, 1981, 9(6):1276～1288

Komlos J, Major R, Tusnady G. An Approximation for Partial Sums of Independent R.V.'s and the Sample D.F., Z. Wahr. Vrew. Geb., 1973, 32: 111～131

Lai T L, Ying Z. Large sample theory of a modified Buckley-James estimator for regression analysis with censored data. Ann Statist 1991, 19:1370～1402

Liu Regina Y C, Ryzin J V. A histogram estimator of the hazard rate with censored data. 1985, 13(2):592～605

Lo S H, SinghK. The product-limit estimator and the bootstrap: some asympototic representations. Probab. Theory Related Fields, 1986, 7:455～465

Lo S H, Wang J L. I. I. D. Representations for the bivariate product limit estimators and the bootstrap versions. Jounal of Mutivariate Analysis, 1989, 28:

211～226

Lynden-Bell D. A method of allowing for known observational selection in small samples applied to 3CR quasar, Monthly Notices Roy. Astronom. Soc., 1971, 155:95～188

Major P, Rejtö L. Strong embedding of the estimator of the distribution function under random censorship. Ann. Statist., 1988, 16(3):1113～1132

Meng X L, Badssiakos Y, Lo S H. Large sample properties for a general estimator of the treatment effect in the two-sample problem with right censoring, Ann. Statist,. 1991, 19:1786～1812

Mielniczuk J. Some asymptotic properties of kernel estimations of a density function in case of censored data, Ann, Statist. 1986, 14:766～773

Miller R, Halpern J. Regression with censored data. Biometika, 1982, 69(3):321～331

Nicall J F, Segal I E. Nonparametric elimination of the observational cutoff bias, Astronom And Astrophys, 1980, 82:3～6

Peterson A V. Expressing the Kaplan-Meier estimator as a function of empirical subsurvival functions. J. Amer. Statist. Assoc., 1977, 72:854～858

Phadia E G, Van Ryzin J. A note on convergence rates for the product-limit estimator. Ann. Statist., 1980, 8:673～678

Priesley M B, Chao M T. Nonparametric function fitting, J. R. Statist. Soc., B, 1972, 34:385～392

Randels R H, Wolfe D A. Introduction To The Theory of Nonparametric Statistics, New York: John Wiley & Sons, 1979

Rao B L S Prakasa. Nonparametric Functionnal Estimation. New York: Academic Press, 1983

Rao B L S Prakasa . Asympototic Theory of Statistical Inference. New York: John Wiley & Sons, 1987. 128～129

Regina Y C, John V R. A Histogram Estimator of the Hazard Rate with Censored Data. Ann. Statist., 1985, 13:592～606

Rice J, Rosenblatt M. Estimation of the Log Survival Function and Hazard Function, Sankhya, Ser A, 1978, 38:60～78

Rosenblatt M. Curve estimates. Ann. Math. Statist. 1971, 42:1815～1842

Rupert G, Miller Jr. Survival Analysis New York: John Wiley & Sons, 1981

Sabine Diehl, Winfried Stute. Kernel Density and Hazard Function Estimation in the Presence of Censoring, J. Mult. Analysis, 1988, 25: 299～310

Serfling R J. Approximation Theorems of Mathematical Statistics. New York: John Wiley & Sons, 1980. 182～183

Shorack G R , Wellner J A. A. s convergence of the Kaplan-Meier estimator, Bulletin. I. M. S. , 1984, 13: 351～365

Shorack G R, Wellner J A. Empirical processes with applications to statistics. New Yowk: John Wiley & Sons, 1986, 293～332

Singh K, Liu R Y, On the validity of the Jack-knife procedure. Scand J. Statist, 1990, 17: 11～21

Susarla H K, Ryzin J V. Regression analysis with randomly right censored data. Ann. Statist. , 1981, 9(6): 1276～1288

Susarla V, Ryzin J V. Nonparametric Bayesian estimation of survival curves from incomplete observation, J. Amer. Statist. Assoc. , 1976, 61: 897～902

Susarla V, Van Ryzin J. Large sample theory for an estimator of themean survivaltime from censored samples. Ann Statist, 1980, 8(5): 1002～1016

Tanner M A, Wong W H. The estimation of the hazard function froom randomly censored data by the kernel methods, Ann. Statist. , 1983, 11: 989～993

Tanner M A. A note on the variable kernel estimator of the Hazard Function from randomly censored data, Ann. Statist. , 1983, 11: 994～998

Tsai W Y , Jewell N P, Wang M C. A note on the product-limit estimator under right censoring and left truncation, Biometrika, 1987, 74(4): 883～886

Tsiatis A A. Estimating regression parameters using linear rank tests for censored data, Ann. Statist. , 1990, 18: 354～372

Tusnay G. An embedding of multivariate empirical distribution Functions, periodical math. , Hunger, 1971, 8: 57～68

Wang J G. A note on the uniform consistency of the Kaplan-Meier estimator. Ann. Statist. 1987, 15: 1313～1316

Wang M C, Jewell N P, Tsai W Y. Asympototic properties of the product limit estimate under random truncation, Ann. Statist. , 1986, 14(4): 1597～1605

Wang Q H. Some Asympototic Behaviours of the Kernel Estimator of Density

Function Based on Censored Data, Chinese Journal of Applied Probability and Statistics, 1994, 10:51～61

Wang Q H. Strong Embedding of Product-Limit Estimator of Bivariate Survival Distribution Function Under Random Censorship, Acta. Mathematica Scientia, 1995, 11(2):123～132

Winter B B, FöldesA, Rejtö L. Glivenko-Cantelli theorems for the product limit estimate, Problem of Control And Information Theory, 1978, 7:213～225

Woodroofe M. Estimating a distribution function with truncated data, Ann. Statist., 1985, 13(1):163～177

Whittle P. Bounds for the moments in linear and quadratic forms in independent random variables. Theory Probab Appl, 1960, 5:302～305

Ying Z L. A large sample study of rank estimation for censored regression data, Ann. Statist., 1993, 21(1):76～99

Zheng Z K. A class of estimators for the parameters in linear regression with censored data, Acta Math. Appl. Sinica., 1987, 3:231～241

Zheng Z K. Strong consistency of nonparametric regression estimates with censored data, Sino-American Statistical Meeting, Beijing, 1987, 627～630

Zheng Z K. Strong uniform consistency for density estimator from randomly censored data, Chin. Ann. Math., B, 1988, 9(2):167～175

Zheng Z K. On the limit points of the Kaplan-Meier estimator. The Third Japan-China Symposium on Statistics. Japan, Tokyo, 1989, 167～170

Zheng Z K . A note on the LIL type result for the product limit estimator, Acta Mathematica Sinica, 1986, 2(2):144～151

Zheng Z K, Limit theorems for the ratio of the Kaplan-Meierestimation or Altshuler estimation to thetrue survival function, Acata Mathematica Sinica, New Series, 1994, 10(4):337～347

Zhou M. Asympototic normality of the "synthetic data" regression estimator for censored survival data, Ann. Statist., 1992, 20:1002～1021

Zhou Y. Nonparametric estimation and inference in random censorship model and random truncation models, Ph. D disertation, Institute of Applied Mathematics, Academic Sinica, Beijing, China, 1993

高集体，赵林城．部分线性回归模型的自适应估计，中国科学，A辑，1992，13:

791～803
高集体. 一类半参数回归模型中估计的相合性,系统科学与数学,1992,12:269～272
洪圣岩.截尾情形下随机窗宽核密度估计,数学学报,1992,35(2):710～718
洪圣岩. 一类半参数回归模型的估计理论,中国科学,A辑,1991,12:1258～1272
胡舒合.固定设计下半参数回归模型估计的强相合性,数学学报,1994,37:395～401
王启华.线性回归模型 L-估计及其光滑 Bootstrap 统计量的一些渐近结构,数学物理学报,1994,14(4):376～384
王启华. 非参数回归函数加权核估计的强相合性,数理统计与应用概率,1992,7(2):43～48
王启华.二维乘积限估计的一些渐近性质.科学通报,1996,41(1):3～5
王启华.基于截尾样本失效率估计的一些大样本性质,科学通报,1994,39(23):2120～2122
王启华.随机截断下半参数回归模型中的相合估计,中国科学,1995,16(8):819～832
郑祖康.截断数据的统计问题,应用概率统计,1988,4(2):17～22